Identification and characterization of Helicobacter pylori O-acetylserine-dependent cystathionine β-synthase, a distinct member of the PLP-II family

Devi, Suneeta; Tarique, Khaja Faisal; Ali, Mohammad Farhan; Abdul Rehman, Syed Arif ; Gourinath, Samudrala

Published in:
Molecular Microbiology

DOI:
10.1111/mmi.14315

Publication date:
2019

Document Version
Peer reviewed version

DR. SAMUDRALA GOURINATH (Orcid ID : 0000-0003-4090-3063)

Article type : Research Article

Identification and characterisation of *Helicobacter pylori* O-acetylserine-dependent cystathionine β-synthase, a distinct member of the PLP-II family

Running title: Structural insights into H. pylori OCBS

Suneeta Devi[1$], Khaja Faisal Tarique[1,2$], Mohammad Farhan Ali[1], Syed Arif Abdul Rehman[1,3] and Samudrala Gourinath[1*]

[1]Structural Biology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi, India
[2]Public Health Research Institute, Rutgers, Newark, NJ, USA
[3]MRC Protein Phosphorylation & Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dundee, UK

[$]These authors contributed equally to work.

[*]To whom correspondence should be addressed:
Prof. S. Gourinath**:** School of Life Sciences, Jawaharlal Nehru University, New Delhi-110067, India
Tel: +91 11 2670 4513, Fax: +91 11 11 2674 2558
E-mail: sgourinath@mail.jnu.ac.in

This is the peer reviewed version of the following article: Devi, S., et al. "Identification and characterisation of Helicobacter pylori O‐acetylserine‐dependent cystathionine β‐synthase, a distinct member of the PLP‐II family", *Molecular Microbiology* (2019), which has been published in final form at https://doi.org/10.1111/mmi.14315. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Summary

O-acetylserine sulfhydrylase (OASS) and cystathionine β-synthase (CBS) are members of the PLP-II family, and involved in L-cysteine production. OASS produces L-cysteine via a *de novo* pathway while CBS participates in the reverse transsulfuration pathway. O-acetylserine-dependent CBS (OCBS) was previously identified as a new member of the PLP-II family, which are predominantly seen in bacteria. The bacterium *Helicobacter pylori* possesses only one OASS (*hp0107*) gene and we showed that the protein coded by this gene actually functions as an OCBS and utilizes L-homocysteine and O-acetylserine (OAS) to produce cystathionine. *Hp*OCBS did not show CBS activity with the substrate L-serine and required OAS exclusively. The *Hp*OCBS structure in complex with methionine showed a closed-cleft state, explaining the initial mode of substrate binding. Sequence and structural analyses showed differences between the active sites of OCBS and CBS, and explain their different substrate preferences. We identified three hydrophobic residues near the active site of OCBS, corresponding to one serine and two tyrosine residues in CBSs. Mutational studies were performed on *Hp*OCBS and *Saccharomyces cerevisiae* CBS. A *Sc*CBS double mutant (Y158F/Y226V) did not display activity with L-serine, indicating indispensability of these polar residues for selecting substrate L-serine, however did show activity with OAS.

Keywords: Cystathionine β-synthase, *Helicobacter pylori*, Hydrogen sulfide, O-acetylserine sulfhydrylase, pyridoxal 5'-phosphate, reverse transsulfuration pathway

Introduction

L-cysteine can be produced via the transsulfuration pathway or a *de novo* pathway depending on the presence of enzymes in the organism. In the transsulfuration pathway, L-methionine is converted into L-homocysteine in multiple steps, with this conversion requiring the action of different enzymes. The L-homocysteine is either converted back to L-methionine via the action of the enzyme L-homocysteine S-methyltransferase or is converted to L-cysteine through an intermediate L-cystathionine via the reverse transsulfuration pathway (Scriver & Kaufman, 2001, Stipanuk, 2004) (Fig. 1). In the reverse transsulfuration pathway, homocysteine reacts with L-serine and is converted to cystathionine by the action of cystathionine β-synthase (CBS), and cystathionine is further acted upon by cystathionine γ-lyase (CGL) to produce L-cysteine. The *de novo* pathway of L-cysteine production from L-serine with O-acetylserine as an intermediate has been shown in bacteria and plants, where serine O-acetyl transferase (SAT) and O-acetylserine sulfhydrylase (OASS) function in tandem (Rabeh & Cook, 2004, Wirtz & Droux, 2005, Kopriva, 2006) (Fig. 1). Recently a few O-acetylserine-dependent CBSs (OCBSs) have been identified and characterised in certain bacteria, have been shown to function in the reverse transsulfuration pathway, specifically being involved in methionine-to-cysteine conversion that takes O-acetylserine from the *de novo* pathway (Devi *et al.*, 2017, Hullo *et al.*, 2007, Matoba *et al.*, 2017).

CBS belongs to the type II fold of PLP-dependent enzymes, and OASS also adopts a similar three-dimensional fold (Alexander *et al.*, 1994, Grishin *et al.*, 1995, Meier *et al.*, 2001, Burkhard *et al.*, 1998). Extensive research has been carried out on *Homo sapiens* CBS (*Hs*CBS), *Droshophila melanogaster* CBS (*Dm*CBS) (Koutmos *et al.*, 2010) and *Saccharomyces cerevisiae* CBS (*Sc*CBS) (Tu *et al.*, 2018). *Hs*CBS and *Dm*CBS each have an extended N-terminal heme-binding region, which probably serves as a redox sensor (Taoka *et al.*, 2002). But neither *Sc*CBS (Tu et al., 2018) nor protozoan CBSs have N-terminal heme-binding region (Nozaki *et al.*, 2001). *Hs*CBS, *Dm*CBS and *Sc*CBS each have a regulatory C-terminal Bateman domain, which is absent in protozoan CBS (Nozaki et al., 2001), but only *Hs*CBS is activated by the binding of the allosteric regulator S-adenosylmethionine (AdoMet) (Finkelstein *et al.*, 1975). All CBSs catalyze the generation of cystathionine via a β-replacement reaction of L-serine with L-homocysteine, along with other alternative β-replacement and β-elimination reactions in which L-cysteine and L-homocysteine are utilized to release H_2S. CBS, especially in the brain, serves as a major source of H_2S, which is involved in modulating the cell function (Chen *et al.*, 2004). Bacterial OCBS also belongs to the PLP type II fold and catalyzes the synthesis of cystathionine via the β-replacement reaction of O-acetylserine and L-homocysteine. OCBS has neither an N-terminal heme-binding region nor a C-terminal regulatory domain (Devi et al., 2017, Matoba et al., 2017). These OCBSs are also capable of catalyzing the production of H_2S efficiently via β-replacement with substrate L-cysteine and L-homocysteine (Devi *et al.*, 2017, Matoba *et al.*, 2017).

Helicobactor pylori is a pathogenic Gram-negative bacterium known to colonize the stomachs of about half of the global human population, largely in developing and resource-poor countries. It afflicts the majority of people with subclinical to progressive clinical gastritis and peptic ulcers (Nesic *et al.*, 2014, Singh *et al.*, 2002). Thus, considering the global etiological magnitude and prominence of the invasiveness of *H. pylori*, this pathogen has been listed as a group 1 carcinogen by the World Health Organisation (WHO) and is considered as the strongest risk factor associated with duodenal and gastric cancers (Peek & Blaser, 2002).

In the *H. pylori* genome, both the SAT and OASS genes have been annotated to belong to the *de novo* pathway, but an overview of the genome of *H. pylori* showed a gene cluster containing the annotated OASS gene (*hp0107*) along with the genes of the cystathionine γ-synthase (*hp0106*) and S-ribosylhomocysteinase (*hp0105*) enzymes of the transsulfuration pathway. This analysis suggested the product of the *hp0107* gene to be also involved in the transsulfuration pathway, where the products of all of these genes are required and thus expressed simultaneously as reported earlier (Doherty *et al.*, 2010). Some recent findings also suggested (Devi et al., 2017, Matoba et al., 2017) that the product of *hp0107* gene could actually be an OAS-dependent CBS (*Hp*OCBS) in *H. pylori*. It should be noted that the transsulfuration pathway has been indicated to be crucial for the viability of *H. pylori* since it grows only in L-methionine-rich media (Mendz & Hazell, 1995, Nedenskov, 1994, Reynolds

& Penn, 1994), and mutations in any one of these genes have been shown to impede the growth of the organism and to result in it becoming a cysteine auxotroph (Doherty et al., 2010).

In the current work, we first carried out an *in silico* study to identify the sub-family to which HP0107 belongs. A phylogenetic tree analysis showed the presence of an OCBS clad separate from OASS and CBS, with the OCBS branch including all annotated OCBSs along with HP0107. This observation indicated HP0107 to be an OCBS (*Hp*OCBS). The *Hp*OCBS protein showed OAS-dependent CBS activity and also catalysed the production of H_2S in the presence of L-cysteine and L-homocysteine. Furthermore, we determined the crystal structure of *Hp*OCBS with bound L-methionine (substrate analog); this analog was observed to be bound near PLP at the active site in one of the protomers of the dimeric structure, explaining the pre-reactive substrate binding state. A comparison of the active sites of OCBS and CBS revealed three key residues playing an important role in substrate selectivity for β-replacement activity to produce cystathionine. Mutation of two of these residues in *Saccharomyces cerevisiae* CBS (*Sc*CBS) resulted in a loss of activity with substrate L-serine and a gain of activity with OAS. Earlier reports on OCBS (Devi et al., 2017, Hullo et al., 2007, Matoba et al., 2017) suggested OCBS to be present only in a few bacteria, but our study revealed the presence of OCBS in a greater variety of bacteria.

Results

In silico study indicates HP0107 is an OCBS

We retrieved many protein sequences annotated either as cysteine synthase or OASS from the NCBI database using PSI-BLAST with *Bacillus anthracis* OCBS as the search template (due to it having previously been shown to be an OCBS (Devi *et al.*, 2017)). Most of these retrieved sequences probably corresponded to OCBS, with the product of the gene *hp0107* of *H. pylori* being one of them. These probable OCBSs derived from the phyla Firmicutes, Chlamydiae, Acidobacteria, Fibrobacter, Actinobacteria and Proteobacteria of Bacteria and Euryochaeota of Archaea (Table S1). A separate list of CBS and OASS proteins from various organisms were also created using PSI-BLAST, with the OASS family containing sequences of OASSs from both bacteria and plants, and the CBS family including sequences from the phyla Arthropoda, Chordata and Nematoda.

Furthermore, a phylogenetic tree was constructed using sequences considered likely to be OCBS, CBS and OASS. A total of 55 probable OCBS sequences (most of them annotated as cysteine synthase in NCBI) clubbed neither with OASS nor with CBS and were instead included in a separate OCBS branch. This separate OCBS branch also contained the product of the gene *hp0107* along with other annotated OCBS sequences as well. This analysis suggested OCBS to differ from OASS and CBS in the PLP-II family of enzymes (Fig. 2) and indicated that the product of the gene *hp0107* belongs to the OCBS subfamily.

Sequence conservation near the active site in OCBS, CBS and OASS

Three multiple sequence alignments were carried out using OCBS, OASS and CBS sequences, respectively, to determine the level of conservation within each of these three types of proteins. The Multalin server was used, which gave a consensus sequence and printed a residue at any particular position when the level of conservation was at least 50% (Fig. S1). The consensus sequence near the active site of the enzyme was divided into five conserved blocks, which showed differences among the sub-families. Inspection of the sequence alignment of conserved block 1, consisting of residues ^{73}TAGNTG^{78} in OCBS sequences (residue numbers according to *Hp*OCBS), indicated TSGNTG to be the corresponding conserved residues in CBS and OASS (Table S2, see also Fig. 3). This sequence in OASS and CBS has been shown to interact using multiple contacts with, respectively, substrate OAS (Schnell *et al.*, 2007) and L-serine (Koutmos *et al.*, 2010) and to also stabilize the α-aminoacrylate intermediate. Inspection of crystal structures indicated this conserved block, also known as the 'asparagine loop', to move upon substrate binding and to close the active site cleft. Conserved block 2, consisting of ^{180}GSGGT^{184} in OCBS, has been indicated to interact with the phosphate moiety of PLP and to anchor the moiety throughout the β-replacement reaction in OCBS (Matoba et al., 2017), OASS (Schnell *et al.*, 2007) and CBS (Koutmos *et al.*, 2010). However, GTGGT was found to be the corresponding sequence conserved in both CBS and OASS. Conserved block 3 ^{221}TEGIGME^{227}, has been predicted to interact with the second substrate (L-homocysteine or HS^- in CBS (Lodha *et al.*, 2009) or OASS, respectively) and to allow this second substrate to participate in a nucleophilic attack on α-aminoacrylate and then release product and regenerate enzyme with an internal aldimine. In the third block, the Met was found to be mostly conserved, but with a Val instead present in *Hp*OCBS. Moreover, these conserved residues were observed to be located near the lip of the active site. Previous analyses of CBS and OCBS crystal structures showed Glu222 of this conserved stretch to form a network of hydrogen bond interactions and to close the active site cleft upon substrate binding (Devi *et al.*, 2017). Sequence alignment revealed IQGIGAG and VEGIGYD to be the block 3 sequences conserved in OASS and CBS, respectively. Tyr in conserved block 3 of CBS has been reported to interact with the substrate L-serine and to play an important role during the β-replacement reaction (Koutmos et al., 2010, Tu et al., 2018). Ser267 of conserved block 4 ^{266}SSSG^{269} of OCBS was observed to interact with the nitrogen of the pyridine ring of PLP; thus, this conserved block may be stabilizing the PLP ring and also holding it during the interconversion between the internal and external aldimines (Schnell *et al.*, 2007, Koutmos *et al.*, 2010). Ser267 and Ser268 were found to be conserved in CBS and OASS as well. Inspection of crystal structures showed conserved block 5 ^{298}RYLSK^{302} to be located at the C-terminal tail and Lys302 of this block to interact with other residues at the lip of active site in OCBS and CBS. And inspection of the sequence alignment indicated RYLST and NYMTK to be the sequences of this block conserved in OASS and CBS, respectively, with the lysine at the fifth position of this block notably conserved in OCBS and CBS. These analyses taken together revealed differences at the sequence level in the conserved blocks in different members of the PLP-II family (Table S2 and Fig. 3).

Kinetics analysis revealed that HP0107 is an OCBS

To confirm that HP0107 is an OCBS, we carried out various biochemical assays. Enzyme was purified to homogeneity to carry out all the enzymatic assays. First, we performed a CBS assay in the presence of substrate L-serine and homocysteine, but it did not show any production of cystathionine. Then OCBS activity was measured to determine whether HP0107 produces L-cystathionine via the β-replacement reaction with OAS and L-homocysteine. HP0107 showed OCBS activity with a k_{cat}/K_m of 12.6 $mM^{-1}\cdot s^{-1}$ for OAS (Table 1 and Fig. 4a), similar to that previously reported from *B. anthracis* (Devi et al., 2017) and *L. plantarum* (Matoba et al., 2017). The HP0107 K_m value for OAS was 2.92 mM, comparable to 1.16 mM and 4.2 mM of OCBS activity of *B. anthracis* (Devi et al., 2017) and *L. plantarum* (Matoba et al., 2017) respectively. Various reported K_m value for L-serine of *Homo sapiens*, *Leishmania major* and *Trypanosoma cruzi* CBS were 1.74 mM, 1.1 mM and 1.0 mM respectively (Chen et al., 2004, Nozaki et al., 2001, Williams *et al.*, 2009), similar to K_m of OCBS activity of HP0107**.** Initially, the OCBS assay was performed using various buffers at different pH values and the highest activity was observed when using HEPES [pH 7.5] buffer. Thus, this buffer was used to carry out all of the reactions. Absence of CBS activity with the substrate L-serine was also observed in the case of other OCBSs, i.e., from *B. anthracis* (Devi *et al.,* 2017), *B. subtilis* (Hullo *et al.,* 2007) and *L. plantarum* (Matoba et al., 2017) and showed production of cystathionine with the substrate O-acetylserine. This result thus indicated HP0107 to be exclusively an OCBS enzyme and to require O-acetylserine instead of L-serine as its first substrate for the reaction (Fig. 5).

HpOCBS releases hydrogen sulfide (H_2S)

CBSs and OCBS produce H_2S in the presence of L-cysteine and L-homocysteine via the β-replacement reaction (Chiku *et al.*, 2009, Devi et al., 2017, Matoba et al., 2017). *Hp*OCBS was also found to produce H_2S with K_m and k_{cat} values for L-homocysteine of 86 µM and 6.05 s^{-1}, respectively (Table 1 and Fig. 4b). The specificity constant of *Hp*OCBS for H_2S generation was observed to be relatively high, when compared with other enzymatic activities of this enzyme. Similarly high efficiency for H_2S production has also been observed in *B. anthracis* (Devi *et al.*, 2017) and *L. plantarum* (Matoba et al., 2017). *Hp*OCBS was also able to produce H_2S using only substrate L-cysteine, which might be producing another product L-lanthionine as suggested by (Matoba et al., 2017) but with very low efficiency (data not shown), compared to the production of H_2S with the β-replacement reaction of cysteine and homocysteine that might be producing another product L-cystathionine (Fig. 5b) as reported (Matoba et al., 2017). The presence of L-homocysteine as a second substrate was found in the current work as well as in previous investigations with *B. anthracis* and *L. plantarum* to greatly enhance the production of H_2S. Inhibition of H_2S production has been shown to make pathogenic bacteria more susceptible to antibiotics (Shatalin *et al.*, 2011). Thus H_2S was concluded to protect the bacteria from oxidative stress, and the higher k_{cat}/K_m indicated its importance for survival of the bacteria.

HpOCBS also shows O-acetylserine sulfhydrylase (OASS) activity

OCBS, like OASS, can consume H_2S to produce L-cysteine (Matoba et al., 2017, Devi et al., 2017). We found HP0107 showing OASS activity, but with a low specificity constant compared to its OCBS and H_2S generation activities. Its k_{cat} was measured to be approximately 0.108 s^{-1} (Table 1 and Fig. 4c), considerably less than the values for OASSs from other organisms (Bonner *et al.*, 2005, Mizuno *et al.*, 2002). OCBS and CBS from other organisms also displayed OASS activity, attributed to their three-dimensional structures and active site residues being similar to those of the OASS proteins.

Three-dimensional structure of HpOCBS reveals its dimeric conformation

The crystal structure of native *Hp*OCBS was determined to a resolution of 1.9 Å (Fig. 6), and refined to an R and R_{free} of 18% and 21%, respectively. Figure 6c shows the quality of the electron density of the PLP molecule and one of the representative regions of the protein. Over 96.4% of the 506 total number of amino acid residues refined in the most favourable region of the Ramachandran plot, and the remaining 3.6% of the residues in the additionally allowed region, i.e., with no residues in the outlier region (Table 2).

The *Hp*OCBS protein crystallized as a dimer with extensive interactions between two protomers (Fig. 6a). The dimeric interface was observed to include both the N- and C-terminal domains. The association between the two monomers was observed to primarily take place through hydrophobic interactions, salt bridges, sulfur-aromatic amino acid interactions (of Phe17 of chain A with Met7 of chain B and chain B Phe17 with chain A Met7) and hydrogen bonds. Together this association resulted in a buried surface area of 3533 $Å^2$, equivalent to 13.4% of the total surface area of the monomer.

Each protomer was observed to consist of two non-contiguous α/β motifs, each having a topology similar to those of other type-II PLP enzymes (Fig. 6b). Inspection of the structure of the N-terminal domain indicated it to consist of residues 1-12 and 40-149, and to include α-helices H1, H2, H3 and H4, β-strands B1, B4, B5, B6 and B7, and 3_{10} helices G1, G3 and G4, with the β-strands mostly parallel to each other and surrounded by helices H2 and H3 from one side and H4 from the other. Inspection of the structure of the C-terminal domain indicated it to consist of residues 13-39 and 150-306, and to include α-helices H5, H6, H7, H8 and H9, β-strands B2, B3, B8, B9, B10 and B11, and 3_{10} helices G2, G5 and G6, with these strands also parallel to each other, and surrounded by helices H6 and H7 from one side and H8 and H9 from the other.

In the *Hp*OCBS crystal structure, PLP cofactor was observed buried in a cleft between the two domains. The cofactor linked to the ε-amino group of Lys44 via a Schiff base linkage to form a lysine-pyridoxal-5`-phosphate (LLP) internal aldimine structure (Fig. 6c). The nitrogen of the pyridine ring (N1) formed a hydrogen bond with the oxygen of Ser267, and the OH (O3) of PLP formed an interaction with Asn76. The phosphate-binding loop or 'G/T loop' was observed to connect β8 and α7 and to be composed of residues ^{180}GSGGT^{184}. These amino acid residues were found to anchor the phosphate moiety by forming a hydrogen

bond network holding the PLP coenzyme firmly (Fig. S2). These residues and the associated hydrogen bonds and hydrophobic interactions near PLP were found to be mostly conserved or similar in other homologous CBS, OCBS and OASS structures.

Methionine complexed with one of the protomers of HpOCBS shows a pre-reactive state

In the *Hp*OCBS crystal structure, methionine (substrate analog) was observed to be bound to one of the protomers of the *Hp*OCBS dimer, at the active site near the PLP. This was, to the best of our knowledge, the first-ever OCBS structure reported with bound methionine along with an internal Schiff base (with PLP covalently linked to the ε-amino group of Lys). In most of the reported substrate-bound structures of this family, PLP is present as an external Schiff base, i.e., with PLP covalently linked to the α-amino group of the substrate or present as an aminoacrylate intermediate (Koutmos et al., 2010, Schnell et al., 2007, Matoba et al., 2017). In two experiments involving OASS, one from *Salmonella typhimurium* (PDB ID: 1D6S) (Burkhard *et al.*, 1999) and the other from *Arabidopsis thaliana* (PDB ID: 1Z7Y) (Bonner et al., 2005), the lysine normally bound to PLP (PLP-linked lysine-LLP) was mutated by the experimenters to Ala, and in these structures PLP makes a covalent bond with methionine (external aldimine). In both of these structures, methionine was not added in the crystallization condition but OASS acquired it during the protein production, indicative of its affinity for methionine. Of all the available OASS structures, only that from *Entamoeba histolytica* has been reported to have methionine at the active site along with the internal aldimine, and this structure was obtained using co-crystallization with the methionine (Raj *et al.*, 2013). Binding constant (K_d) values of methionine were reported to be 78 ± 9 μM and 0.70 ± 0.06 mM, respectively, for the OASSs of *S. typhimurium* and *A. thaliana*, with the latter value comparable to the 0.252 ± 0.055 mM value measured here for *Hp*OCBS (Fig. S3). Based in part on this result, *Hp*OCBS may be predicted to have a high affinity for methionine and trapped methionine in the pre-reactive state from the media during protein production. The presence of methionine in the pre-reactive state in only one of the protomers (of the dimeric *Hp*OCBS) and without it covalently bound to the PLP to form an external aldimine were attributed to the suboptimal concentration of the methionine in our experiments. However, a high K_d value of 1.5 ± 0.4 mM was reported for methionine with *E. histolytica* OASS, indicative of the inability of this protein to easily take up methionine from media, and thus the need to have added it externally during the crystallization (Chinthalapudi *et al.*, 2008).

Superposition of the two protomers of the *Hp*OCBS dimer revealed an approximately 0.78 Å root-mean-square deviation (RMSD) of backbone α-carbon atoms, similar to that found for L-serine-bound *D. melanogaster* CBS (Koutmos *et al.*, 2010). Methionine binding led to a closing of the active site as a result of the movement of mobile loop 222-230 of the C-terminal domain and a nearly 10° bending of the N-terminal domain (Fig. 7c), making the two domains approach each other, and specifically decreasing the distance between the domains by 3.2 Å (Fig. 7c). Here, in the presence of methionine, Glu222 (conserved in OCBSs (Devi *et al.*, 2017) and CBSs) formed interactions with Lys302 of the C-terminal tail and with the backbone amide of Ser101 of the N-terminal domain (Fig. 7d).

Methionine in *Hp*OCBS was observed to form a hydrogen bond with the side chain of Lys104. A mutation of the corresponding lysine in yCBS (K112) was carried out previously by Lodha *et. al*, 2009 to determine its importance in L-homocysteine binding, but surprisingly it yielded an increase in the K_m of L-serine instead of L-homocysteine, indicating that this particular Lys binds to L-serine instead of L-homocysteine (Lodha *et al.*, 2009). This conclusion was also consistent with the above-indicated observation of an interaction between methionine and Lys104 in the present methionine-bound *Hp*OCBS structure. The backbone nitrogen atom of Asn76 of the ^{73}TAGNT^{77} 'asparagine loop' also made a hydrogen bond with the carboxyl group of methionine, while other residues of this loop formed hydrophobic interactions with methionine (Fig. 7a and b) and apparently promoted closing of the active site (Burkhard *et al.*, 1999). Ser has been found to be conserved instead of Ala74 in the asparagine loops of CBSs, and mutation of Ala to Ser in OCBS yielded an increase in H_2S generation (Devi et al., 2017, Matoba et al., 2017), indicating the involvement of the Ser residue in anchoring the substrate.

Gln146, conserved across OCBS, CBS and OASS sequences, was observed to interact with methionine in *Hp*OCBS to form an internal Schiff base. A Gln157Ala mutation in yeast CBS abolished β-replacement reaction and this result confirmed the role of this residue in substrate positioning and in the subsequent reaction mechanism (Aitken & Kirsch, 2004). In the current *Hp*OCBS crystal structure, Phe147and Gly223 also interacts with the methionine (Fig. 7a and b). Also note that the residue corresponding to *Hp*OCBS Phe147 is a Tyr in CBS, and a Tyr158Phe mutation in yeast CBS was previously shown to yield a 3-fold decrease in k_{cat}/K_m for L-serine in the β-replacement reaction (Aitken & Kirsch, 2004). The hydroxyl group of this Tyr was reported to form a hydrogen bond network with the β-OH group of the substrate L-serine and other residues at the active site of CBS and hence to facilitate the β-replacement reaction (Koutmos et al., 2010).

Structural comparison of HpOCBS with other homologous enzymes of the PLP-II family

We compared the *Hp*OCBS structure with the available structures of CBS, OCBS and OASS of the PLP-II family. CBS was found to have N- and C-terminal domains with three-dimensional structures similar to those of *Hp*OCBS. In addition to these domains, CBSs from higher organisms also have C-terminal extended regulatory and N-terminal heme-binding domains (Omura *et al.*, 1984, Finkelstein *et al.*, 1975, Koracevic & Djordjevic, 1977). In mammals, CBS is activated by the binding of AdoMet to the regulatory domain (Prudova *et al.*, 2006, Finkelstein *et al.*, 1975, Koracevic & Djordjevic, 1977), while CBS proteins from different classes remain constitutively active and do not require AdoMet for the activation (Koutmos et al., 2010, Gimenez-Mascarell *et al.*, 2017). Four CBS structures have been reported, from *H. sapiens* (Meier et al., 2001), *D. melanogaster* (Koutmos et al., 2010), *Apis mellifera* (Gimenez-Mascarell et al., 2017) and *S. cerevisiae* (Tu et al., 2018), and 3-dimentional structural superposition of these CBSs with *Hp*OCBS showed RMSD values of approximately 2.0 Å for about 300 aligned Cα atoms of the core catalytic domain (i.e., N- and C-terminal domains) (Table 3 and Fig. 8b). Sequence comparison of CBSs with *Hp*OCBS showed only 23 to 28 % sequence identity values (Table 3), with these low values due to the presence of extra residues at both the N- and C-termini.

To date, only two OCBS structures have been published, from *B. anthracis* (PDB ID: 5XW3) (Devi *et al.,* 2017) and *L. plantarum* (PDB ID: 5B1H) (Matoba et al., 2017), and rigid body superpositions of these structures with the *Hp*OCBS structure using the FATCAT server gave RMSD values of 3.0 Å and 1.15 Å, respectively. The very high RMSD between *Hp*OCBS and *B. anthracis* OCBS was due to the large angle between the N- and C-terminal domains in the latter. On comparing the methionine-bound *H. pylori* OCBS structure (PDB ID: 6AHI) with the methionine-bound *L. plantarum* OCBS structure (PDB ID: 5B1I), we found the methionine in the *H. pylori* structure to be present as a pre-reactive state and not linked to PLP (Fig. 9b). This methionine is expected to adopt a conformation similar to the methionine of *L. plantarum* OCBS after making a Schiff base linkage with PLP.

Structural superposition of *Hp*OCBS with known OASSs yielded RMSD values of 1.8 to 2.0 Å, indicating their structures to be highly similar; these superpositions also yielded very good structural alignments of their active site residues (Fig. 8a). Sequence comparisons of *Hp*OCBS with OASSs showed only 33 to 40 % identities but with the OCBSs showed nearly 53 to 58 % identities (Table 3). These results clearly justified the placement of *Hp*OCBS in a separate OCBS clad in the phylogenetic tree.

Mode of OAS and L-homocysteine binding at active site

To date, no CBS or OCBS structure with bound L-homocysteine has been published. Therefore, to gain an understanding of the modes of interactions of OAS and L-homocysteine at the active site for the β-replacement reaction, L-homocysteine was docked *in silico* in the closed conformation of the methionine-bound *Hp*OCBS structure to identify the residues involved in the binding of L-homocysteine. Here, L-homocysteine bound between the asparagine loop and the loop containing residues 222 to 227 of the C-terminal domain (Fig. S4a). The docking produced hydrogen-bond interactions of the backbone of Gly223 and Gly225 with the α-carboxylate and α-amino group of L-homocysteine (Fig. S4b and c).These same residues formed interactions with a distal portion of L-cystathionine docked in yeast CBS, with this distal portion of L-cystathionine corresponding to the L-homocysteine moiety (Lodha *et al.*, 2009). Residues Glu222, Ile224, Glu227 and His219 from the C-terminal domain and Phe100 and Ile124 from the N-terminal domain formed hydrophobic contacts with L-homocysteine (Fig. S4b and c). Methionine was observed to be located in between PLP and L-homocysteine, more towards the N-terminal domain below the asparagine loop (Fig. S4a).

Mutational studies confirm the critical active site residue differences between conventional CBS and OCBS

Conventional CBS prefers L-serine as its first substrate to carry out β-replacement reaction, while OCBSs exclusively require OAS for the reaction and do not show any functionality in the presence of the substrate L-serine (Devi *et al.,* 2017), with this different behaviour towards OAS and L-serine perhaps due to the difference between certain residues at the active site of CBS and those of OCBS (Table S2 and Fig. 9a). Inspection of the *Hp*OCBS and CBS structures specifically showed three residues located near the active site, at (*Hp*OCBS) positions 74, 147 and 226, to differ between them —identified, respectively, as Ala, Phe, and

Val in *Hp*OCBS, but Ser, Tyr and Tyr in CBS. The hydroxyl groups of the Ser and Tyr residues may provide necessary hydrogen bonding to the leaving hydroxyl (β-OH) group of the L-serine substrate and facilitate its β-elimination (Koutmos *et al.*, 2010), while absence of these residues in OCBS, may be making it incapable of utilizing L-serine for the reactions and thus requiring activated L-serine, i.e., OAS, for its activity.

To confirm the identity of the residues crucial for the reaction mechanism, a series of mutations of *Hp*OCBS were made (listed in Table S3). The β-replacement activity of each mutation was checked with the substrate L-serine or OAS in the presence of 5 mM of homocysteine. Specifically three structure-guided mutations, A74S, F147Y, and V226Y, were made to *Hp*OCBS. In addition to these single-residue-substitution mutants, one double-residue-substitution mutant (F147Y/V226Y) and one triple-residue-substitution mutant (A74S/F147Y/V226Y) were also generated. All five of these mutants were generated with the expectation that they would show activity with the substrate L-serine, with this expectation due to the interactions made between the β-OH of L-serine-bound CBS structures (Koutmos et al., 2010) (Tu et al., 2018) and the homologous residues. However, none of the mutants showed β-replacement activity with the L-serine (Fig. S5), even though there was no structural change upon mutation (Fig. S6). The F147Y mutant showed activity with OAS but with less catalytic efficiency than did wild type (WT) *Hp*OCBS. Surprisingly, the V226Y variant did not show any β-replacement activity with substrate OAS. We checked OCBS activity with a higher enzyme concentration (1.48 μM) of the *Hp*OCBS V226Y mutant together with 5mM each of OAS and L-homocysteine; only slight activity was observed in this case. Also, neither the double mutant (F147Y/V226Y) nor the triple mutant (A74S/F147Y/V226Y) of *Hp*OCBS showed any activity with substrate OAS even at the high enzyme concentration, again despite the lack of any considerable structural differences between the mutant and WT *Hp*OCBS structures (Fig. S6).

To verify our hypothesis on the importance of these residues for substrate binding occurring via the β-replacement reaction mechanism, mutation was made to *S. cerevisiae* CBS (*Sc*CBS) at the corresponding homologous residues listed in Table S3. *Sc*CBS was chosen due to its lack of an N-terminal iron-binding region and due to the constitutive activity of its catalytic core domain, which has been shown to not require AdoMet for activation. Also, *Sc*CBS has been shown to be stable for nearly 6-7 days at 4°C. *Sc*CBS showed k_{cat} and K_m values for CBS activity to be 55.6 s^{-1} and 2.26 mM respectively. The discrepancy in the k_{cat} and K_m values of CBS activity of *Sc*CBS with the earlier report may have been due to the difference in the method employed for detecting cystathionine (Table 4); however the k_{cat}/K_m of *Sc*CBS was determined to be 24.68 $mM^{-1}s^{-1}$, comparable to the previously reported value (Lodha *et al.*, 2010). The double mutant (Y158F/Y248V) generated in *Sc*CBS was tested for cystathionine generation in the presence of substrate L-serine and OAS, individually. The *Sc*CBS double mutant (Y158F/Y248V) did not show any CBS activity with L-serine but showed enhanced activity in the presence of substrate OAS, with a k_{cat}/K_m of about 18.2 $mM^{-1}s^{-1}$ (Table 4), compared to the 12.6 $mM^{-1}s^{-1}$ value for the WT *Sc*CBS OCBS activity (Fig. S5). Based on these results, we hypothesized that Y248 and Y158 are indispensable for CBS activity with L-serine not only in yeast CBS but in all CBSs, since these tyrosine residues are conserved across this sub-family. However, the OCBS activity displayed by the *Sc*CBS

double mutant (Y158F/Y248V) showed a higher specificity constant than did that displayed by WT *Sc*CBS (Table 4 and Fig. S5), indicating the favorability of the active site architecture for OAS, hence validating the importance of these residues for OAS binding.

The absence of CBS activity with substrate L-serine by the triple mutant (A74S/F147Y/V226Y) and double mutant (F147Y/V226Y) of *Hp*OCBS could be due to the different orientation adopted by Val226 in *Hp*OCBS than that adopted by the homologous Tyr248 in *Sc*CBS. In the *Hp*OCBS crystal structure, Val226 faces away from the active site towards the C-terminal domain and its movement seems to be restricted due to the presence of Glu227 which interacts with the residues of N-terminal domain in a closed conformation upon substrate binding (methionine-bound protomer of *Hp*OCBS) (Matoba et al., 2017). Thus even upon mutation of Val226 to Tyr in *Hp*OCBS, this residue would still presumably be facing away from the active site and hence unable to contact the substrate L-serine.

Discussion

The reverse transsulfuration pathway is well characterised in higher organisms such as *H. sapiens* (Ereno-Orbea *et al.*, 2013, Ereno-Orbea *et al.*, 2014, McCorvie *et al.*, 2014), *S. cerevisiae* (Jhee *et al.*, 2000), *A. mellifera* (Gimenez-Mascarell *et al.*, 2017) and *D. melanogaster* (Koutmos et al., 2010), where it is the sole source of L-cysteine production from methionine. In bacteria, the reverse transsulfuration pathway is also known as the 'methionine-to-cysteine conversion pathway'. CBS and CGL have been named, respectively, as MccA and MccB enzymes of the methionine-to-cysteine conversion pathway. This pathway is not universally present in bacteria but study shows some bacteria possess enzymes to carry out methionine-to-cysteine conversion. In *Lactobacillus casei*, the expression levels of *cbs* and *cgl* genes decrease in cysteine-containing media; these genes are present in a reading frame along with the gene for serine O-acetyl transferase (Bogicevic *et al.*, 2012). Various *Lactobacillus paracasei* strains possess the highly mobile gene cluster *cysK-ctl-cysE* responsible for the conversion of methionine to cysteine in dairy products, which are highly rich in methionine compared to cysteine (Wuthrich *et al.*, 2018). Similarly, *Klebsiella pneumoniae* uses its *mtcBC* operon to produce cystathionine β-synthase and cystathionine γ-lyase (Seiflein & Lawrence, 2006) for the production of cysteine from methionine. *Mycobacterium tuberculosis* has a bifunctional gene (*rv1079),*which allows this organism to carry out both cystathionine γ-lyase and cystathionine γ-synthase activities and use methionine as a sole sulfur source, and hence demonstrate a functional reverse transsulfuration pathway (Wheeler *et al.*, 2005). *Bacillus subtilis* possesses CBS, which shows activity with OAS and not with serine and can use methionine as a sole source of sulfur since it also has the *cgl* gene (Hullo et al., 2007). Recent investigations have also shown the importance of CBS for the survivability and pathogenicity of bacterial pathogens under oxidative stress (Shatalin et al., 2011, Devi et al., 2017, Matoba et al., 2017).

The present study along with earlier studies showed the presence of the OCBS gene in the genomes of some bacteria, and showed these bacterial OCBSs to differ from the CBSs found in higher organisms in their domain architecture, length, regulation and most importantly substrate requirements. These studies showed bacterial OCBSs to function in the presence of

OAS instead of L-serine to produce L-cystathionine, to not require AdoMet for its activation, and to also be devoid of a heme-binding domain and C-terminal regulatory domain (Devi et al., 2017, Matoba et al., 2017). OCBS was shown to be widely found in bacteria, and to share a high similarity with OASS, and in fact has been misannotated as either cysK or cysteine synthase (OASS). OASS has been shown to be involved in different activities in addition to its role in cysteine production in the *de novo* pathway, such as an antibacterial nuclease toxin produced by *Escherichia coli* (Campanini *et al.*, 2015) and in transcription repression in *B. subtilis* (Tanous *et al.*, 2008). OCBS has apparently maintained its integrity at the sequence level, showing no mixing either with OASS and CBS in the phylogenetic tree (Fig. 2). The *H. pylori hp0107* gene is annotated as an OASS in the NCBI database but the protein encoded by this gene was found in the current work to show OCBS activity and to also display O-acetylserine sulfhydrylase activity, albeit at a relatively low level, a feature observed for other OCBSs as well (Matoba et al., 2017, Devi et al., 2017). Based on the phylogenetic analysis and analysis of the *Hp*OCBS structure with bound methionine (substrate analog) at the active site, we concluded the present *hp0107* gene to actually code for an OCBS.

An in depth structural comparison of OCBS with conventional CBS, which uses L-serine, revealed possible key residues that determine the respective substrate preferences of these two proteins. Analysis of the active sites was performed by using L-serine-bound *S. cerevisiae* CBS (PDB ID: 6C2Q) (Tu et al., 2018) and methionine-bound *Hp*OCBS. We identified three residues that may play important roles in selecting L-serine as the substrate for the reaction performed by CBS. Inspection of the *Sc*CBS structure showed Ser82 and Tyr248 forming hydrogen bonds with the hydroxyl group of the substrate L-serine, and this substrate covalently linked to PLP (forming an external aldimine). A network of polar interactions of Tyr158 and Tyr248 would appear to also stabilize the reaction intermediate (Fig. 9a). Inspection of conventional CBS also showed its active site cavity to be polar, due to the presence of two tyrosine residues and one serine residue, with these three residues ultimately making polar interactions with the hydroxyl group of the L-serine substrate and participating in its activation and dehydroxylation. Analysis of the OCBS structures showed them to contain a conserved Ala74 (residue number according to *H. pylori* OCBS) instead of Ser82 in the 'TSGNT' loop, while showing Tyr158 replaced by Phe147 and Val226 (or Met or any other hydrophobic amino acid) at the homologous position of Tyr248 in the 'EGIGY' loop (Fig. 9a). These residues make the active site of OCBS hydrophobic and unable to participate in the activation of L-serine. Interestingly, the *Sc*CBS double mutant containing Phe and Val instead of Tyr (Y158F/Y248V) behaved like an OCBS and showed β-replacement activity only with substrate OAS. Perhaps CBS evolved from bacterial OCBS to take L-serine instead of OAS as its substrate, with evidence for this speculation including the observed lack of the enzyme to produce OAS in higher organisms and the measured retention of ancestral OCBS activity in the *Sc*CBS mutants.

Nevertheless, to complete the reverse transsulfuration pathway, another enzyme known as cystathionine γ-lyase (CGL) is required (Banerjee *et al.*, 2003, Miles & Kraus, 2004). In *H. pylori*, the *hp0106* gene (annotated as cystathionine γ-synthase) is in the same gene cluster as that encoding OCBS but has been predicted to function as a CGL (Doherty *et al.*, 2010). Our

investigation also revealed the ability of HP0106 to efficiently convert cystathionine to cysteine (data not shown). Thus functional OCBS and CGL are present in *H. pylori* and a complete reverse transsulfuration pathway is functional in this pathogen.

Experimental procedures

In silico analysis

Various OCBS protein sequences were retrieved using PSI-Blast with *B. anthracis* O-acetylserine-dependent CBS (*Ba*OCBS) as the search template (Devi *et al.*, 2017). Furthermore, a few resultant sequences were used as search templates to retrieve additional distantly related OCBSs present in various phyla of bacteria. We used here only a few protein sequences from any particular genus. This means that there are many such proteins present in different species of the same genus. CBS proteins present in different organisms were retrieved using *H. sapiens* CBS (Ereno-Orbea et al., 2014) as a template in PSI-Blast. However, annotated CBSs present in a few bacteria were retrieved from the NCBI database manually. Similarly, OASS sequences from bacteria and plants were also extracted using *E. histolytica* OASS (PDB ID: 2PQM) as a template sequence in PSI-Blast (Chinthalapudi *et al.*, 2008). Finally, a total of 154 sequences were used for the construction of the phylogenetic tree. Initially a multiple sequence alignment was done using Clustal Omega with all of the sequences. The aligned sequences were then used as input into MEGA 5.1 software and a phylogenetic tree was generated using maximum likelihood based on the default parameters with 250 bootstraps (Tamura *et al.*, 2011).

Furthermore, multiple sequence alignment was done using the MULTALIN tool to obtain a consensus sequence for the members of the CBS, OASS and OCBS families (Corpet, 1988). All sequences from one particular family were input into the server for the multiple sequence alignment, which used the BLOSUM62 comparison table. The output file gave the aligned sequences along with the consensus sequence, which included residues conserved in more than 50% of the sequences.

Amplification, cloning and site-directed mutagenesis of HP0107 and ScCBS

The *hp0107* gene was amplified by performing the polymerase chain reaction (PCR) on the genomic DNA of *H. pylori* 26695 as a template using gene specific primers (Table 5) and Phusion High-Fidelity DNA Polymerase (Thermo Scientific^{TM}). The amplicon of the 921 bp gene was first purified using a PCR purification kit (Qiagen, Germany) and digested with NcoI and XhoI at 37°C for five hours. The pET28b expression vector was also digested with NcoI and XhoI, and the digested vector was purified using a gel extraction kit (Qiagen, Germany). The purified *hp0107* gene was ligated, using T4 DNA ligase, into the pET28b expression vector at the NcoI and XhoI sites with a C-terminal $(His)_6$ tag. The positive construct was selected and screened using a Luria-Bertani (LB) agar plate supplemented with kanamycin (50 $\mu g \cdot ml^{-1}$) and colony PCR, respectively. The site-directed mutagenesis was performed on a pET28b-*hp0107* plasmid using mutagenic primers for the corresponding mutations listed in Table 5. In brief, pET28b-*hp0107* was used as a template for PCR to

generate variants. Following the amplification of the whole plasmid, parental methylated DNA was digested using DpnI enzyme. The digested mixture was directly transformed in the DH5α strain of *E. coli*. All of the intended mutations were confirmed by carrying out whole gene sequencing.

Similarly, amplification of the catalytic core region of the *Sccbs* gene was done by PCR on genomic DNA of *S. cerevisiae* with the primers listed in Table 5, followed by cloning of *Sccbs* in pET28b at the NheI and XhoI sites. Furthermore, site-directed mutagenesis was performed on the *Sccbs* gene as described above.

Overexpression and purification of HP0107 and ScCBS

Plasmids of pET28b-*hp0107* and pET28b-*Sccbs* were transformed in the BL21 (DE3) strain of *E. coli* for the protein expression. The colonies were picked from LB plates containing kanamycin and inoculated into 50 ml of LB media for overnight growth. One percent of this culture was inoculated into one litre of secondary culture containing 50 $\mu g \cdot ml^{-1}$ of kanamycin. The culture was incubated at 37°C until the OD_{600} reached 0.6 to 0.8. Gene expression was initiated with the addition of isopropyl β-D-thiogalactoside (IPTG) to a final concentration of 250 µM, the temperature was lowered to 16°C after induction, and the culture was allowed to grow for another 12 h. Cells were harvested at 4°C by subjecting them to centrifugation at 7500 g for 5 minutes and then resuspending the pellet in buffer A containing 50 mM Tris [pH7.5], 150 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF), 0.1% Triton X-100, 5 mM imidazole, 0.01 $mg \cdot ml^{-1}$ lysozyme and 5 mM β-mercaptoethanol (βMe). Then, cells were sonicated and centrifuged at 25 000 g for 1 h. The cleared supernatant was applied onto Ni-NTA Sepharose column that had previously been equilibrated with 50 ml buffer B (50 mM Tris [pH 7.5], 150 mM NaCl, 5 mM βMe and 5 mM imidazole). The protein was eluted in buffer C containing 50 mM Tris [pH7.5], 150 mM NaCl, 300 mM imidazole, 5% glycerol and 5 mM βMe.

For further purification and removal of imidazole, the protein was concentrated using a Centricon filter with 30 kDa cut-off and loaded onto a HiLoad 16/60 Superdex 200 column (GE Healthcare) and eluted at a flow rate of 1 $ml \cdot min^{-1}$. The column that had previously been equilibrated with running buffer D containing 10 mM Tris [pH 7.5], 150 mM NaCl, 5% glycerol and 5 mM βMe. The purity of the desired protein was checked using 12% SDS-PAGE and purified fractions were pooled and concentrated using a Centricon filter (cut-off 30 kDa). Nearly 95 to 99% pure protein was obtained from gel-filtration chromatography (GFC) and purified protein was used to perform all of the experiments.

Purification of the *hp0107* mutant proteins were done using a method similar to that used for WT HP0107 (*Hp*OCBS), and an SDS-PAGE analysis of the purified proteins obtained after GFC showed 90 to 99% pure protein (Fig. S7). The expression levels of the *Hp*OCBS A74S and *Hp*OCBS F147Y single mutants were similar to that of WT *Hp*OCBS, each approximately 8 to 10 mg of protein from one litre of culture, while a mass of only 3 to 4 mg of protein was obtained from the same volume of culture for the *Hp*OCBS V226Y mutant

and the *Hp*OCBS double (F147Y/V226Y) mutant (Fig. S8). The expression level of the *Hp*OCBS triple mutant (A74S/F147Y/V226Y) was 6 to 7 mg from one litre of culture (Fig. S8), better than that of the *Hp*OCBS double mutant but poorer than that of WT *Hp*OCBS. Furthermore, the *Hp*OCBS V226Y mutant and *Hp*OCBS double (F147Y/V226Y) mutant were observed to be less stable that the WT *Hp*OCBS and tended to precipitate when kept at 4°C for more than 3 to 4 days, but the *Hp*OCBS A74S and *Hp*OCBS F147Y mutants were found to be as stable as *Hp*OCBS. The *Hp*OCBS triple mutant (A74S/F147Y/V226Y) was observed to be more stable than the *Hp*OCBS double mutant (F147Y/V226Y) and remained in solution for 5 to 6 days at 4°C. The stability of this triple mutant may have been provided in part by the hydrogen bonding made between Ser at position 74 and the PLP.

WT *Sc*CBS and its *Sc*CBS double (Y158F/Y428V) mutant were expressed at high levels, approximately 25 to 30 mg of protein from 500 mL of culture. No significant difference was observed between the expression and stability of the *Sc*CBS double mutant and those of WT *Sc*CBS (Fig. S9).

Biochemical characterisation

O-acetylserine-dependent CBS (OCBS) assay

An O-acetylserine-dependent CBS assay was conducted using a colorimetric method as described earlier for CBS (Kashiwamata & Greenberg, 1970). An OCBS enzyme assay was performed in a final volume of 200 µl with final concentrations of 0.367 µM for *Hp*OCBS, 0.25 mM for pyridoxal 5`-phosphate and various concentrations for OAS in 50 mM HEPES [pH 7.5] buffer, while keeping the L-homocysteine concentration constant at 5 mM, to determine the K_m for OAS. In each experiment, the reaction mixture was incubated for 5 minutes at 37°C, and then the reaction was stopped with the addition of ninhydrin reagent directly into the reaction mixture. The reaction solution was heated for 5 minutes on a boiling water bath and then incubated for 3 minutes on ice, after which it was left to develop colour for 20 minutes at 25°C. The absorbance of the solution was measured at a wavelength of 451 nm and from this absorbance value was subtracted the absorbance of an enzyme-free blank. Initially, this assay was performed with various buffers of different pH values (50 mM each of sodium acetate [pH 4.0 to 5.0], Bis-Tris [pH 5.5 to 6.5], HEPES [pH 7.0 to 8.1] and Tris-HCl [pH 8.5 to 8.8]) to determine the optimal pH for the OCBS activity. In initial assays, the concentrations of L-homocysteine and OAS were kept constant at 1 mM and 5 mM, respectively. The amount of cystathionine produced in a reaction was calculated from the standard curve of the cystathionine (50 to 5000 µM). All results shown represent mean ± SD of three (n=3) independent experiments. Data obtained from all of the above assays were fitted by the Michaelis-Menten equation $v = (V_{max}[OAS])/(K_m+[OAS])$ using GraphPad Prism 5 to calculate the K_m and V_{max} values of the reactions.

Cystathionine β-synthase (CBS) assay

CBS assay was also carried out with substrate L-serine and L-homocysteine as described above for the OCBS assay. All of the reactions with *Hp*OCBS and its variants were

performed in 50 mM HEPES [pH 7.5] with a final enzyme concentration of 0.367 µM. Reactions were performed with a constant amount of homocysteine (5 mM) and various concentrations of freshly prepared L-serine along with 0.25 mM of PLP. Reaction mixtures were incubated at 37 °C for 5 minutes and stopped with the addition of ninhydrin.

The CBS and OCBS assays performed with the *Sc*CBS protein used 0.30 µM of the *Sc*CBS enzyme, 0.25 mM pyridoxal 5`-phosphate and 5 mM of L-homocysteine in 50 mM of Tris [pH 8.6] buffer as done earlier (Lodha et al., 2009).

Data obtained from the CBS assay were fitted by the Michaelis-Menten equation $v = (V_{max}$ [L-Ser])/(K_m+[L-Ser]) using GraphPad Prism 5 to calculate the K_m and V_{max} values of the reactions while the data for the OCBS assay were fitted by the equation $v = (V_{max}$[OAS])/(K_m+[OAS]) using GraphPad Prism 5 to calculate the K_m and V_{max} values of these reactions. Each result shown is the mean ± SD of three (n=3) independent experiments.

H_2S production assay

OCBSs and CBSs produce H_2S via the β-replacement reaction with L-cysteine and L-homocysteine as substrates (Devi et al., 2017, Banerjee et al., 2003, Matoba et al., 2017). The production of H_2S catalyzed by *Hp*OCBS was monitored by measuring the absorption (at λ=390 nm) due to lead sulfide formed from the reaction of H_2S with lead acetate. A reaction mixture of 400 µl containing 0.4 mM lead acetate, various concentrations of L-homocysteine, and a fixed concentration of L-cysteine (1 mM) in 50 mM HEPES [pH 7.5] was preincubated at 25°C. The reaction was started by adding 10 µg (final concentration 0.367µM) of enzyme, and the absorption was monitored at a wavelength of 390 nm for 2 minutes. A molar extinction coefficient of 5,500 $M^{-1}cm^{-1}$ was used to determine the amount of lead sulfide produced (Chiku *et al.*, 2009). Each result shown is the mean ± SD of three (n=3) independent experiments.

The Michaelis-Menten equation $v = (V_{max}$ [L-homocysteine])/(K_m+[L-homocysteine]) was used to fit the data obtained from the H_2S production assay and hence calculate the K_m and V_{max} values of the reaction.

O-acetylserine sulfhydrylase (OASS) assay

The OASS activity was determined as described previously (Kumar *et al.*, 2011). Briefly, the reaction mixtures (final volume 400 µL) contained 1.58 µM (final concentration) of enzyme, 100 µM of 5-thio-2-nitrobenzoate TNB (alternative substrate) and various concentrations of OAS (6.75 to 250 µM) in 50 mM HEPES [pH 7.5] buffer. The reaction was initiated with the addition of OAS to the mixture of enzyme and TNB, and readings were taken immediately at a wavelength 412 nm for 3 minutes. The extinction coefficient of the TNB (A_{412}; ε=13,600 $M^{-1}cm^{-1}$) was used to calculate the rates of the reactions (Bonner et al., 2005). All of the experiments were done at 25°C and the results shown each represented a mean ± SD of three (n=3) independent experiments.

Data obtained from the OASS assay were fitted by the Michaelis-Menten equation $v = (V_{max}[OAS])/(K_m+[OAS])$ using GraphPad Prism 5 to calculate the K_m and V_{max} values of the reactions.

Methionine binding

Binding of methionine to OASS results in the formation of an external aldimine with PLP. Methionine binding to *Hp*OCBS was measured using a spectrofluorometer to monitor the changes in fluorescence spectra of this complex during the course of the binding. Spectra were taken from 450 nm to 600 nm, upon excitation at wavelength 412 nm, while keeping the slit width at 5 nm. Titration of *Hp*OCBS (100 µg) with methionine (0 to 700 µM) was performed in 50 mM HEPES [pH 7.5] buffer and an appropriate concentration of methionine was employed to achieve the saturation curve. All of the experiments were performed at 25°C. The maximum change in the fluorescence was observed at 516 nm. The binding constant (K_d) of methionine was calculated by fitting to the data a reversible two-state model of binding using the equation $\Delta F = (\Delta F_{max}[M]/ K_d + [M])$, where ΔF represents the change in fluorescence at a particular wavelength upon addition of methionine at concentration [M] which is corrected for dilution (Bonner et al., 2005, Banerjee *et al.*, 2011). ΔF_{max} represents the maximum change in the fluorescence at that particular wavelength. The binding constant (K_d) of methionine for *Hp*OCBS was calculated to be 0.252 ± 0.055 mM (Fig. S3).

Crystallization and X-ray diffraction data collection

The purified *Hp*OCBS protein was concentrated up to 8 $mg \cdot ml^{-1}$in a solution containing 10 mM Tris [pH 7.5] with 150 mM NaCl, 5% glycerol and 5 mM βMe. After continuous optimization using the hanging-drop vapour-diffusion method, crystals appeared in 3 to 4 days at 16°C using a mixture of 15% PEG 20000, 15% PEG MME 550, and 0.03 M NPS (sodium nitrate, disodium hydrogen phosphate and ammonium sulfate) at pH 6.5 (0.1 M MES/imidazole) as the precipitant, this condition was similar to the Morpheus crystallization screen condition number C1(Gorrec, 2009).

The needle-shaped crystals that formed were carefully picked up, mounted in cryoloops, and flash frozen in liquid nitrogen at 100 K. These flash-frozen crystals diffracted X-rays (λ=0.978Å) to a resolution of 1.9 Å at ESRF BM14, Grenoble, France (Table 2). The data were indexed and processed using HKL2000 (Otwinowski & Minor, 1997). These crystals were shown to belong to the space group $P22_12_1$ with two *Hp*OCBS protomers per asymmetric unit.

Structure solution and refinement

The *Hp*OCBS structure was determined by carrying out molecular replacement using the structure of O-acetylserine sulfhydrylase from *Arabidopsis thaliana* (AtCS) (PDB ID: 1Z7W, 43% sequence identity). This AtCS structure and the aligned sequences of *Hp*OCBS and AtCS were input into CHAINSAW (Schwarzenbacher *et al.*, 2004) to produce the specific search model used for molecular replacement, which was then carried out using Phaser

(McCoy *et al.*, 2007) of CCP4 (Winn *et al.*, 2011). This procedure gave a single solution with a high translation function Z-score (TFZ score) and log-likelihood gain (LLG). An initial cycle of restrained refinement using this solution and Refmac5 (Murshudov *et al.*, 1997) resulted in a drop of the R-factor to 30%. The model was then submitted to ARP/wARP for auto-building, which successfully built 93% of the side chains of the model into good quality electron density. The remaining parts of the polypeptide were built manually with COOT, and later the PLP and water molecules were added according to the electron density.

Multiple rounds of manual adjustments of the model in COOT (Emsley & Cowtan, 2004) and refinement with Refmac5 (Murshudov *et al.*, 1997) were carried out. This process resulted in a final model for *Hp*OCBS consisting of 4680 protein atoms and 142 water molecules, and that fit well into the electron density and yielded good refinement statistics with an R and R_{free} of 18% and 21%, respectively (Table 2). One of the protomers adopted a closed cleft conformation and had extra density that appeared to be methionine with 50% occupancy. The structure factors and coordinates of this model have been deposited in the Protein Data Bank with the PDB ID: 6AHI.

Circular dichroism spectropolarimetry

Far-UV CD spectra were recorded using a Chirascan^{TM} Plus CD spectrometer from Applied Photophysics (Surrey, UK) equipped with a Peltier element, in a 1.0 mm quartz cuvette between wavelengths of 260 nm and 200 nm at a rate of 1 nm per step. Spectra of WT *Hp*OCBS and mutant *Hp*OCBS were taken at a protein concentration of 0.15 mg·mL^{-1}. Each spectrum represented the average of 10 scans taken at 20ºC. Analysis of these spectra indicated similar secondary structures for all the *Hp*OCBS mutants and WT *Hp*OCBS (Fig. S6a).

Thermal unfolding of the protein at a concentration of 0.15 mg·mL^{-1} was monitored by acquiring the CD responses at 222 nm with increasing temperature at a rate of 1°C·min^{-1} from 25 to 90°C. *Hp*OCBS A74S and the *Hp*OCBS F147Y mutant showed comparable midpoints of thermal denaturation (T_m), each at about 72.96 ± 0.71 and 72.61 ± 1.22°C respectively (Table S4) and hence indicated to be slightly more stable than *Hp*OCBS which showed T_m of 70.55 ± 0.31°C, while *Hp*OCBS V226Y was shown to be less stable, with a T_m of 63.55 ± 0.15°C. The *Hp*OCBS double mutant (F147Y/V226Y) and *Hp*OCBS triple mutant (A74S/F147Y/V226Y) were also found to be less stable, with T_m of only about 66.06 ± 1.08 and 55.12 ± 1.44°C respectively (Fig. S6 and Table S4). These melting temperature were calculated as described elsewhere (Ubaid-Ullah *et al.*, 2014). All CD experiments were performed in 20 mM of phosphate buffer [pH 7.5].

Similarly, we also acquired far-UV CD scans and monitored the thermal unfolding of *Sc*CBS and its mutant, each at concentration of 0.15 mg·mL^{-1}. The T_m values of *Sc*CBS and its mutant were each determined to be about 59.2°C as reported earlier (Tu et al., 2018) (Fig. S10 and Table S4).

Docking of L-homocysteine at the active site of HpOCBS structure

The Met-bound *Hp*OCBS structure was used for the docking of L-homocysteine at the active site. First, we prepared the *Hp*OCBS protein structure for the docking using the Schrodinger Protein Preparation Wizard program with the default parameters. This program was used to add missing hydrogen atoms, remove water molecules beyond 5 Å of HET groups, determine the protonated state for histidine residues, and highlight the missing residues. Further energy minimization and optimization of the protein structure was done using an OPLS2005 force field with a maximum RMSD of 0.30 Å for atom displacement to terminate the minimization. A ligand-binding site or ligand receptor grid was created in the energy-minimized structure of *Hp*OCBS near the PLP. The ligand receptor grid covered all of the expected residues of the active site from the N- and C-terminal domains of *Hp*OCBS. The L-homocysteine substrate was prepared using Schrodinger Ligprep Wizard before docking; here, hydrogen atoms were added, followed by energy minimization. Finally, docking of substrate in the *Hp*OCBS receptor grid was performed using the program GLIDE with the extra precision (XP) module.

Accession number

The accession number (PDB ID) for *Hp*OCBS structure reported in this paper is **6AHI**

Acknowledgements

We thank the Department of Science and Technology (DST), Department of Biotechnology (DBT), Council of Scientific and Industrial Research (CSIR), Government of India for funding this project. SD and MFA are thankful to the Indian Council of Medical Research (ICMR) for a Senior Research Fellowship and DST-SERB, Government of India, New Delhi for a National Postdoctoral Fellowship, respectively. We thank DST-FIST, DBT-BUILDER and DST-PURSE for funding the central instrumentation facility and for extending institutional funding. We thank Prof. Suman Dhar, JNU, New Delhi for providing genomic DNA of *Helicobacter pylori* 26695. We also thank the staffs of BM14, ESRF for helping us in the collection of high-resolution X-ray data.

Declarations of interest

All authors declare no conflict of interest.

Authors' contributions

SG, SD and KFT conceived of the aim of the work and designed the experiments. KFT, SD and SAAR solved the *Hp*OCBS structure. SD and MFA performed all experiments. SD, KFT and MFA wrote the manuscript taking input from SG and SAAR. All of the authors reviewed the manuscript.

REFERENCES

Aitken, S. M. & J. F. Kirsch, (2004) Role of active-site residues Thr81, Ser82, Thr85, Gln157, and Tyr158 in yeast cystathionine beta-synthase catalysis and reaction specificity. *Biochemistry* **43**: 1963-1971.

Alexander, F. W., E. Sandmeier, P. K. Mehta & P. Christen, (1994) Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families. *Eur J Biochem* **219**: 953-960.

Banerjee, R., R. Evande, O. Kabil, S. Ojha & S. Taoka, (2003) Reaction mechanism and regulation of cystathionine beta-synthase. *Biochim Biophys Acta* **1647**: 30-35.

Banerjee, S., M. K. Ekka & S. Kumaran, (2011) Comparative thermodynamic studies on substrate and product binding of O-acetylserine sulfhydrylase reveals two different ligand recognition modes. *BMC Biochem* **12**: 31.

Bogicevic, B., S. Irmler, R. Portmann, L. Meile & H. Berthoud, (2012) Characterization of the cysK2-ctl1-cysE2 gene cluster involved in sulfur metabolism in Lactobacillus casei. *Int J Food Microbiol* **152**: 211-219.

Bonner, E. R., R. E. Cahoon, S. M. Knapke & J. M. Jez, (2005) Molecular basis of cysteine biosynthesis in plants: structural and functional analysis of O-acetylserine sulfhydrylase from Arabidopsis thaliana. *J Biol Chem* **280**: 38803-38813.

Burkhard, P., G. S. Rao, E. Hohenester, K. D. Schnackerz, P. F. Cook & J. N. Jansonius, (1998) Three-dimensional structure of O-acetylserine sulfhydrylase from Salmonella typhimurium. *J Mol Biol* **283**: 121-133.

Burkhard, P., C. H. Tai, J. N. Jansonius & P. F. Cook, (2000) Identification of an allosteric anion-binding site on O-acetylserine sulfhydrylase: structure of the enzyme with chloride bound. *J Mol Biol* **303**: 279-286.

Burkhard, P., C. H. Tai, C. M. Ristroph, P. F. Cook & J. N. Jansonius, (1999) Ligand binding induces a large conformational change in O-acetylserine sulfhydrylase from Salmonella typhimurium. *J Mol Biol* **291**: 941-953.

Campanini, B., R. Benoni, S. Bettati, C. M. Beck, C. S. Hayes & A. Mozzarelli, (2015) Moonlighting O-acetylserine sulfhydrylase: New functions for an old protein. *Biochim Biophys Acta* **1854**: 1184-1193.

Chattopadhyay, A., M. Meier, S. Ivaninskii, P. Burkhard, F. Speroni, B. Campanini, S. Bettati, A. Mozzarelli, W. M. Rabeh, L. Li & P. F. Cook, (2007) Structure, mechanism, and conformational dynamics of O-acetylserine sulfhydrylase from Salmonella typhimurium: comparison of A and B isozymes. *Biochemistry* **46**: 8315-8330.

Chen, X., K. H. Jhee & W. D. Kruger, (2004) Production of the neuromodulator H2S by cystathionine beta-synthase via the condensation of cysteine and homocysteine. *J Biol Chem* **279**: 52082-52086.

Chiku, T., D. Padovani, W. Zhu, S. Singh, V. Vitvitsky & R. Banerjee, (2009) H2S biogenesis by human cystathionine gamma-lyase leads to the novel sulfur metabolites lanthionine and homolanthionine and is responsive to the grade of hyperhomocysteinemia. *J Biol Chem* **284**: 11601-11612.

Chinthalapudi, K., M. Kumar, S. Kumar, S. Jain, N. Alam & S. Gourinath, (2008) Crystal structure of native O-acetyl-serine sulfhydrylase from Entamoeba histolytica and its complex with cysteine: structural evidence for cysteine binding and lack of interactions with serine acetyl transferase. *Proteins* **72**: 1222-1232.

Corpet, F., (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**: 10881-10890.

Crooks, G. E., G. Hon, J. M. Chandonia & S. E. Brenner, (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.

Devi, S., S. A. Abdul Rehman, K. F. Tarique & S. Gourinath, (2017) Structural characterization and functional analysis of cystathionine beta-synthase: an enzyme involved in the reverse transsulfuration pathway of Bacillus anthracis. *FEBS J* **284**: 3862-3880.

Dharavath, S., I. Raj & S. Gourinath, (2017) Structure-based mutational studies of O-acetylserine sulfhydrylase reveal the reason for the loss of cysteine synthase complex formation in Brucella abortus. *Biochem J* **474**: 1221-1239.

Doherty, N. C., F. Shen, N. M. Halliday, D. A. Barrett, K. R. Hardie, K. Winzer & J. C. Atherton, (2010) In Helicobacter pylori, LuxS is a key enzyme in cysteine provision through a reverse transsulfuration pathway. *J Bacteriol* **192**: 1184-1192.

Emsley, P. & K. Cowtan, (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**: 2126-2132.

Ereno-Orbea, J., T. Majtan, I. Oyenarte, J. P. Kraus & L. A. Martinez-Cruz, (2013) Structural basis of regulation and oligomerization of human cystathionine beta-synthase, the central enzyme of transsulfuration. *Proc Natl Acad Sci U S A* **110**: E3790-3799.

Ereno-Orbea, J., T. Majtan, I. Oyenarte, J. P. Kraus & L. A. Martinez-Cruz, (2014) Structural insight into the molecular mechanism of allosteric activation of human cystathionine beta-synthase by S-adenosylmethionine. *Proc Natl Acad Sci U S A* **111**: E3845-3852.

Finkelstein, J. D., W. E. Kyle, J. L. Martin & A. M. Pick, (1975) Activation of cystathionine synthase by adenosylmethionine and adenosylethionine. *Biochem Biophys Res Commun* **66**: 81-87.

Gimenez-Mascarell, P., T. Majtan, I. Oyenarte, J. Ereno-Orbea, J. Majtan, J. Klaudiny, J. P. Kraus & L. A. Martinez-Cruz, (2017) Crystal structure of cystathionine beta-synthase from honeybee Apis mellifera. *J Struct Biol*.

Gorrec, F., (2009) The MORPHEUS protein crystallization screen. *Journal of applied crystallography* **42**: 1035-1042.

Grishin, N. V., M. A. Phillips & E. J. Goldsmith, (1995) Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci* **4**: 1291-1304.

Hullo, M. F., S. Auger, O. Soutourina, O. Barzu, M. Yvon, A. Danchin & I. Martin-Verstraete, (2007) Conversion of methionine to cysteine in Bacillus subtilis and its regulation. *J Bacteriol* **189**: 187-197.

Jhee, K. H., P. McPhie & E. W. Miles, (2000) Domain architecture of the heme-independent yeast cystathionine beta-synthase provides insights into mechanisms of catalysis and regulation. *Biochemistry* **39**: 10548-10556.

Kashiwamata, S. & D. M. Greenberg, (1970) Studies on cystathionine synthase of rat liver. Properties of the highly purified enzyme. *Biochim Biophys Acta* **212**: 488-500.

Kopriva, S., (2006) Regulation of sulfate assimilation in Arabidopsis and beyond. *Ann Bot* **97**: 479-495.

Koracevic, D. & V. Djordjevic, (1977) Effect of trypsin, S-adenosylmethionine and ethionine on L-serine sulfhydrase activity. *Experientia* **33**: 1010-1011.

Koutmos, M., O. Kabil, J. L. Smith & R. Banerjee, (2010) Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine {beta}-synthase. *Proc Natl Acad Sci U S A* **107**: 20958-20963.

Kumar, S., I. Raj, I. Nagpal, N. Subbarao & S. Gourinath, (2011) Structural and biochemical studies of serine acetyltransferase reveal why the parasite Entamoeba histolytica cannot form a cysteine synthase complex. *J Biol Chem* **286**: 12533-12541.

Lodha, P. H., E. M. Hopwood, A. L. Manders & S. M. Aitken, (2010) Residue N84 of yeast cystathionine beta-synthase is a determinant of reaction specificity. *Biochim Biophys Acta* **1804**: 1424-1431.

Lodha, P. H., H. Shadnia, C. M. Woodhouse, J. S. Wright & S. M. Aitken, (2009) Investigation of residues Lys112, Glu136, His138, Gly247, Tyr248, and Asp249 in the active site of yeast cystathionine beta-synthase. *Biochem Cell Biol* **87**: 531-540.

Matoba, Y., T. Yoshida, H. Izuhara-Kihara, M. Noda & M. Sugiyama, (2017) Crystallographic and mutational analyses of cystathionine beta-synthase in the H2 S-synthetic gene cluster in Lactobacillus plantarum. *Protein Sci* **26**: 763-783.

McCorvie, T. J., J. Kopec, S. J. Hyung, F. Fitzpatrick, X. Feng, D. Termine, C. Strain-Damerell, M. Vollmar, J. Fleming, J. M. Janz, C. Bulawa & W. W. Yue, (2014) Inter-domain communication of human cystathionine beta-synthase: structural basis of S-adenosyl-L-methionine activation. *J Biol Chem* **289**: 36018-36030.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni & R. J. Read, (2007) Phaser crystallographic software. *Journal of applied crystallography* **40**: 658-674.

Meier, M., M. Janosik, V. Kery, J. P. Kraus & P. Burkhard, (2001) Structure of human cystathionine beta-synthase: a unique pyridoxal 5'-phosphate-dependent heme protein. *EMBO J* **20**: 3910-3916.

Mendz, G. L. & S. L. Hazell, (1995) Aminoacid utilization by Helicobacter pylori. *Int J Biochem Cell Biol***27**: 1085-1093.

Miles, E. W. & J. P. Kraus, (2004) Cystathionine beta-synthase: structure, function, regulation, and location of homocystinuria-causing mutations. *J Biol Chem* **279**: 29871-29874.

Mizuno, Y., Y. Miyashita, S. Yamagata, T. Iwama & T. Akamatsu, (2002) Cysteine synthase of an extremely thermophilic bacterium, Thermus thermophilus HB8. *Biosci Biotechnol Biochem* **66**: 549-557.

Murshudov, G. N., A. A. Vagin & E. J. Dodson, (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* **53**: 240-255.

Nedenskov, P., (1994) Nutritional requirements for growth of Helicobacter pylori. *Appl Environ Microbiol* **60**: 3450-3453.

Nesic, D., L. Buti, X. Lu & C. E. Stebbins, (2014) Structure of the Helicobacter pylori CagA oncoprotein bound to the human tumor suppressor ASPP2. *Proc Natl Acad Sci U S A* **111**: 1562-1567.

Nozaki, T., Y. Shigeta, Y. Saito-Nakano, M. Imada & W. D. Kruger, (2001) Characterization of transsulfuration and cysteine biosynthetic pathways in the protozoan hemoflagellate, Trypanosoma cruzi. Isolation and molecular characterization of cystathionine beta-synthase and serine acetyltransferase from Trypanosoma. *J Biol Chem* **276**: 6516-6523.

Omura, T., H. Sadano, T. Hasegawa, Y. Yoshida & S. Kominami, (1984) Hemoprotein H-450 identified as a form of cytochrome P-450 having an endogenous ligand at the 6th coordination position of the heme. *The Journal of Biochemistry* **96**: 1491-1500.

Otwinowski, Z. & W. Minor, (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307-326.

Peek, R. M., Jr. & M. J. Blaser, (2002) Helicobacter pylori and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer* **2**: 28-37.

Prudova, A., Z. Bauman, A. Braun, V. Vitvitsky, S. C. Lu & R. Banerjee, (2006) S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. *Proc Natl Acad Sci U S A* **103**: 6489-6494.

Rabeh, W. M. & P. F. Cook, (2004) Structure and mechanism of O-acetylserine sulfhydrylase. *J Biol Chem* **279**: 26803-26806.

Raj, I., S. Kumar & S. Gourinath, (2012) The narrow active-site cleft of O-acetylserine sulfhydrylase from Leishmania donovani allows complex formation with serine acetyltransferases with a range of C-terminal sequences. *Acta Crystallogr D Biol Crystallogr* **68**: 909-919.

Raj, I., M. Mazumder & S. Gourinath, (2013) Molecular basis of ligand recognition by OASS from E. histolytica: insights from structural and molecular dynamics simulation studies. *Biochim Biophys Acta* **1830**: 4573-4583.

Reynolds, D. J. & C. W. Penn, (1994) Characteristics of Helicobacter pylori growth in a defined medium and determination of its amino acid requirements. *Microbiology* **140 (Pt 10)**: 2649-2656.

Schnell, R., W. Oehlmann, M. Singh & G. Schneider, (2007) Structural insights into catalysis and inhibition of O-acetylserine sulfhydrylase from Mycobacterium tuberculosis. Crystal structures of the enzyme alpha-aminoacrylate intermediate and an enzyme-inhibitor complex. *J Biol Chem* **282**: 23473-23481.

Schwarzenbacher, R., A. Godzik, S. K. Grzechnik & L. Jaroszewski, (2004) The importance of alignment accuracy for molecular replacement. *Acta crystallographica. Section D, Biological crystallography* **60**: 1229-1236.

Scriver, C. R. & S. Kaufman, (2001) Disorders of Transulfuration. *Scriver, CR, Kaufman, S., Eisensmith, E., Woo SLC, Vogelstein, B. Childs, B.(eds) The Metabolic and Molecular Bases of Inherited Disease, 8th ed. McGraw-Hill, New York.*

Seiflein, T. A. & J. G. Lawrence, (2006) Two transsulfurylation pathways in Klebsiella pneumoniae. *J Bacteriol* **188**: 5762-5774.

Shatalin, K., E. Shatalina, A. Mironov & E. Nudler, (2011) H2S: a universal defense against antibiotics in bacteria. *Science* **334**: 986-990.

Singh, V., B. Trikha, C. K. Nain, K. Singh & K. Vaiphei, (2002) Epidemiology of Helicobacter pylori and peptic ulcer in India. *J Gastroenterol Hepatol* **17**: 659-665.

Stipanuk, M. H., (2004) Sulfur amino acid metabolism: pathways for production and removal of homocysteine and cysteine. *Annu Rev Nutr* **24**: 539-577.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei & S. Kumar, (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731-2739.

Tanous, C., O. Soutourina, B. Raynal, M. F. Hullo, P. Mervelet, A. M. Gilles, P. Noirot, A. Danchin, P. England & I. Martin-Verstraete, (2008) The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in Bacillus subtilis. *J Biol Chem* **283**: 35551-35560.

Taoka, S., B. W. Lepore, O. Kabil, S. Ojha, D. Ringe & R. Banerjee, (2002) Human cystathionine beta-synthase is a heme sensor protein. Evidence that the redox sensor is heme and not the vicinal cysteines in the CXXC motif seen in the crystal structure of the truncated enzyme. *Biochemistry* **41**: 10454-10461.

Tu, Y., C. A. Kreinbring, M. Hill, C. Liu, G. A. Petsko, C. D. McCune, D. B. Berkowitz, D. Liu & D. Ringe, (2018) Crystal Structures of Cystathionine beta-Synthase from Saccharomyces cerevisiae: One Enzymatic Step at a Time. *Biochemistry* **57**: 3134-3145.

Ubaid-Ullah, S., M. A. Haque, S. Zaidi, M. I. Hassan, A. Islam, J. K. Batra, T. P. Singh & F. Ahmad, (2014) Effect of sequential deletion of extra N-terminal residues on the structure and stability of yeast iso-1-cytochrome-c. *J Biomol Struct Dyn* **32**: 2005-2016.

Wheeler, P. R., N. G. Coldham, L. Keating, S. V. Gordon, E. E. Wooff, T. Parish & R. G. Hewinson, (2005) Functional demonstration of reverse transsulfuration in the Mycobacterium tuberculosis complex reveals that methionine is the preferred sulfur source for pathogenic Mycobacteria. *J Biol Chem* **280**: 8069-8078.

Williams, R. A., G. D. Westrop & G. H. Coombs, (2009) Two pathways for cysteine biosynthesis in Leishmania major. *Biochem J* **420**: 451-462.

Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin & K. S. Wilson, (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**: 235-242.

Wirtz, M. & M. Droux, (2005) Synthesis of the sulfur amino acids: cysteine and methionine. *Photosynth Res* **86**: 345-362.

Wuthrich, D., S. Irmler, H. Berthoud, B. Guggenbuhl, E. Eugster & R. Bruggmann, (2018) Conversion of Methionine to Cysteine in Lactobacillus paracasei Depends on the Highly Mobile cysK-ctl-cysE Gene Cluster. *Front Microbiol* **9**: 2415.

TABLES

Table 1. Various kinetics parameters of the reactions catalyzed by *Hp*OCBS

	Substrate	**V_{max} (μM s^{-1})**	**K_m (mM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($mM^{-1}s^{-1}$)**
OCBS activity[a]	O-acetylserine	13.51 ± 0.52	2.92 ± 0.28	36.8 ± 1.4	12.6 ± 1.31
CBS activity[a]	L-serine	NA[¶]	NA	NA	NA
H_2S-production activity[b]	L-homocysteine	2.22 ± 0.05	0.086 ± 0.007	6.05 ± 0.13	70.33 ± 5.98
OASS activity[c]	O-acetylserine	0.17 ± 0.006	0.025 ± 0.003	0.108 ± .004	4.42 ± 0.17

[¶] NA- No detectable activity. [a]OCBS and CBS kinetics experiments were carried out in 50 mM HEPES [pH 7.5] buffer at 37°C. [b] H_2S-production activity experiments were carried out in 50 mM HEPES [pH 7.5] buffer at 25°C. [c] OASS activity was determined in 50mM HEPES [pH 7.5] buffer at 25°C. Each result shown is the mean ± SD of three (n=3) independent experiments.

Table 2. *Hp*OCBS crystallographic statistics

Data statistics	*Hp*OCBS
Crystallographic data	
Diffraction source	ESRF beamline BM14
Wavelength (Å)	0.978
Space group	$P22_12_1$
a, *b*, *c* (Å)	71.4, 82.7, 96.1
α, β, γ (°)	90, 90, 90
Resolution range (Å)	50–1.9 (1.93–1.9)[¶]
$R_{r.i.m.}$	0.121 (0.88)
R-meas	0.117 (0.862)
Completeness (%)	100 (100)
Redundancy	16.3 (15.8)
$\langle I/\sigma(I)\rangle$	47 (4.4)
Overall *B* factor from Wilson plot (Å^2)	24.4
CC1/2	0.99 (0.97)
No. of molecules per asymmetric unit	2
Refinement	
Total No. of reflections	741356
No. of unique reflections	45368
Final R_{cryst}	0.183
Final R_{free}	0.207
No of protein/water atoms	4667/132
R.m.s. deviations	
Bonds (Å)	0.010
Angles (°)	1.469
Average *B* factors (Å^2)	
Protein	37.5
Water	33.4
Ramachandran plot	
Favoured (%)	96.4
Allowed (%)	3.6

[¶] Values for the highest-resolution shell are shown in parentheses.

Table 3. Comparisons of the structure and sequence of *Hp*OCBS with those available for OCBS, OASS and CBS. Table shows RMSDs of Cα atoms with number of equivalent residues and percentage sequence identity values.

	PDB ID	RMSD (Å)	Equivalent residues	Sequence identity (%)
OCBS	5B1H (Matoba et al., 2017)	1.15	301	58.1
	5XW3 (Devi et al., 2017)	3.06	257	54.9
OASS	1OAS (Burkhard *et al.*, 2000)	1.98	299	37.5
	1VE1[¶]	1.69	298	37.2
	1Z7W (Bonner et al., 2005)	1.80	303	40.2
	2JC3 (Chattopadhyay *et al.*, 2007)	1.88	289	34.9
	2PQM (Chinthalapudi et al., 2008)	1.87	304	35.4
	2Q3B (Schnell et al., 2007)	1.96	297	40.8
	3SPX (Raj *et al.*, 2012)	1.85	302	35.2
	5JIS (Dharavath *et al.*, 2017)	2.01	302	33.9
CBS	1JBQ (Meier et al., 2001)	1.93	301	28.8
	3PC2 (Koutmos et al., 2010)	1.98	305	23.2
	5OHX (Gimenez-Mascarell et al., 2017)	2.00	305	27.1
	6C2H(Tu et al., 2018)	2.10	304	33.8

[¶]O-acetylserine sulfhydrylase from *Thermus thermophilus HB8*

Table 4. Kinetics parameters of *Hp*OCBS variants, and of *Sc*CBS WT and its variants.

Protein	Variant	Substrate	V_{max} (µM s^{-1})	K_m (mM)	k_{cat}(s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
*Hp*OCBS[a]	A74S	OAS	8.43 ± 0.56	1.78 ± 0.39	22.96 ± 1.53	12.89 ± 2.96
		L-serine	NA[¶]	NA	NA	NA
	F147Y	OAS	5.77 ± 0.32	1.57 ± 0.29	15.72 ± 0.87	10.04 ±1.92
		L-serine	NA	NA	NA	NA
	V226Y	OAS	NA	NA	NA	NA
		L-serine	NA	NA	NA	NA
	F147Y/ V226Y	OAS	NA	NA	NA	NA
		L-serine	NA	NA	NA	NA
	A74S/F 147Y/V 226Y	OAS	NA	NA	NA	NA
		L-serine	NA	NA	NA	NA
*Sc*CBS[b]	*Sc*CBS WT	OAS	6.89 ± 0.67	1.82 ± 0.43	22.50 ± 2.20	12.35 ± 3.13
		Lserine	17.03 ± 1.15	2.26 ± 0.37	55.65 ± 3.75	24.68 ± 4.41
	Y158F/ Y248V	OAS	14.67 ± 1.06	2.69 ± 0.53	47.94 ± 3.47	17.85 ± 9.47
		L-serine	NA	NA	NA	NA

[¶] NA-No detectable activity. [a]OCBS and CBS kinetics experiments were carried out in 50 mM HEPES [pH 7.5] buffer at 37 °C for WT *Hp*OCBS and mutant proteins. [b]OCBS and CBS kinetics experiments of *Sc*CBS and its mutant were carried out in 50 mM Tris-HCl [pH 8.6] buffer at 37 °C. Each result shown is the mean ± SD of three (n=3) independent experiments.

Table 5. Primers sequences used in present study.

Primer	Sequences
*Hp*OCBS_Fp[¶]	5'-CATGCCATGGCAATGATGATTATCACCACAATG-3'
*Hp*OCBS_Rp[¶]	5'-CGGCTCGAGTAAATAAATACCTTTTGAGAGATAA-3'
*Hp*OCBS_A74S_Fp	5'-CATCATTGAGCCTACCAGCGGCAATACCGGCATCG-3'
*Hp*OCBS_A74S_Rp	5'-CGATGCCGGTATTGCCGCTGGTAGGCTCAATGATG-3'
*Hp*OCBS_F147Y_Fp	5'-AGCTATTTGCCCTTACAATATGAAAACCCTGATAATCCCG-3'
*Hp*OCBS_F147Y_Rp	5'-CGGGATTATCAGGGTTTTCATATTGTAAGGGCAAATAGCT-3'
*Hp*OCBS_V226Y_Fp	5'-AGATTGAGGGCATTGGCTATGAGTTCATCCCTCCTTT-3'
*Hp*OCBS_V226Y_Rp	5'-AAAGGAGGGATGAACTCATAGCCAATGCCCTCAATCT-3'
*Sc*CBS_Fp[a]	5'-CTAGCTAGCATGACTAAATCTGAGCAGCAAG-3'
*Sc*CBS_Rp[a]	5'-CGGCTCGAGTCACAGCTTTGAAGAGTCAAAACG-3'
*Sc*CBS_Y158F_Fp	5'-GTTATACTTGACCAATTTAACAATATGATGAAC-3'
*Sc*CBS_Y158F_Rp	5'-GTTCATCATATTGTTAAATTGGTCAAGTATAAC-3'
*Sc*CBS_Y248V_Fp	5'-AGTTGAGGGTATTGGTGTGGATTTTGTTCCTCAGG-3'
*Sc*CBS_Y248V_Rp	5'-CCTGAGGAACAAAATCCACACCAATACCCTCAACT-3'

[¶] underlined sequences represent the restriction sites, while Fp and Rp stand for forward and reverse primers, respectively.

FIGURE LEGENDS

Figure 1. Reverse transsulfuration pathway and *de novo* pathway in *Helicobacter pylori.* In the reverse transsulfuration pathway, L-homocysteine reacts with L-serine in the presence of CBS to produce the intermediate L-cystathionine, which is further converted to L-cysteine by the action of the enzyme cystathionine γ-lyase (CGL) (Scriver & Kaufman, 2001, Stipanuk, 2004) (shown with green arrow). *H. pylori* can produce L-cysteine utilizing L-methionine via the reverse transsulfuration pathway. The organism includes OCBS (HP0107) and can produce L-cystathionine from L-homocysteine and OAS (an intermediate in the *de novo* pathway) (shown with red arrow). HP0106 is annotated as cystathionine γ-synthase but predicted to function as CGL in *H. pylori* (Doherty et al., 2010). *H. pylori* also has *de novo* pathway enzymes, namely serine O-acetyltransferase (SAT) and O-acetylserine sulfhydrylase OASS (OCBS in *H. pylori)*, and can produce L-cysteine. In *H. pylori,* OASS was shown in the current work to function as an OCBS.

Figure 2. Phylogenetic study of OASS, OCBS and CBS of the PLP-II family. A phylogenetic analysis was carried out with 154 sequences using the maximum likelihood method with 250 bootstraps of the Mega 5.1 software. This analysis yielded a phylogenetic tree made up of three different clads, namely OCBS (red), OASS(green) and CBS (blue), with the OCBS clad containing well-annotated OCBSs from various bacteria, including *Hp*OCBS (PDB ID: 6AHI). Red stars indicate experimentally verified OCBS proteins while blue stars indicate sequences predicted, but not yet shown experimentally, to be OCBS in whole genome sequencing of the corresponding bacterial genome.

Figure 3. WebLogo generated from multiple sequence alignment of OCBSs (55 sequences), CBSs (52 sequences) and OASSs (47 sequences). Residue numbering is according to *Hp*OCBS for the OCBS alignment, *H. sapiens* CBS for the CBS alignment and OASS of *M. tuberculosis* for the OASS alignment. Consensus sequences near the active site were divided into 5 conserved blocks and the residues in CBS and OASS differing from those in OCBS are shown with red stars. Height of an individual amino acid in Weblogo indicates level of conservation at the particular portion (Crooks *et al.*, 2004).

Figure 4. **Biochemical characterisation of *Hp*OCBS.** (a) The OAS-dependent CBS activity of *Hp*OCBS was characterised with 5 mM L-homocysteine and various concentrations of OAS in 50 mM of HEPES [pH 7.5] to determine the K_m and V_{max} values of the reaction. *Hp*OCBS also catalyzes the production of H_2S from the reaction of L-cysteine with L-homocysteine. (b) Continuous monitoring of hydrogen sulfide production in the presence of 0.4 mM lead acetate upon reaction of L-cysteine (1 mM) with increasing concentrations of L-homocysteine. (c) *Hp*OCBS was also found to display OASS activity, which was measured with various concentrations of OAS in 50 mM HEPES [pH 7.5] in the presence of 100 μM of TNB. Each result shown is the mean ± SD of three (n=3) independent experiments.

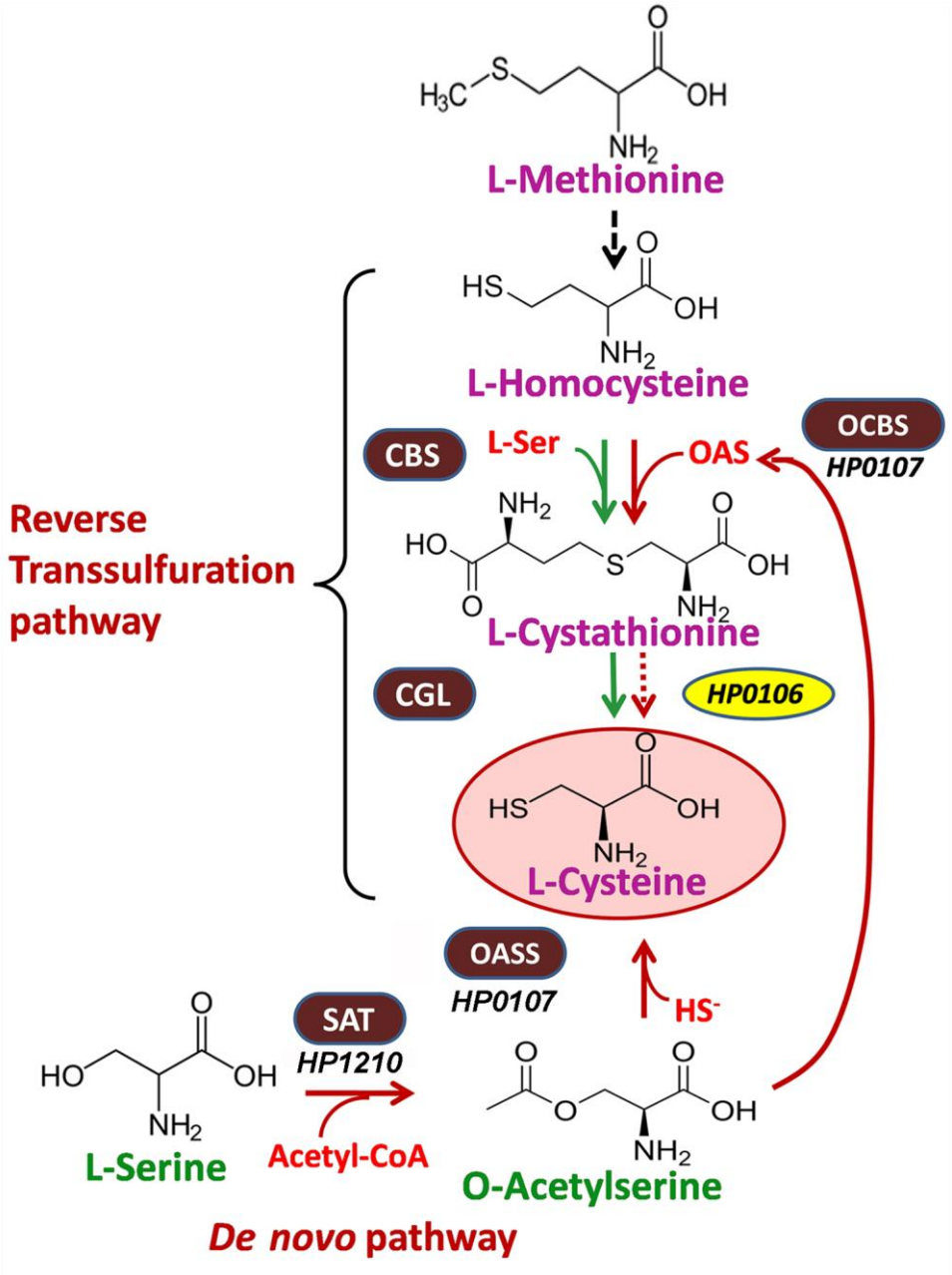
Figure 5. Reactions catalyzed by *Hp*OCBS. 1) *Hp*OCBS shows OCBS activity with substrates OAS and L-homocysteine and catalyzes the production of L-cystathionine. 2) *Hp*OCBS also catalyzes the production of H_2S with substrates L-cysteine and homocysteine. 3) OASS activity is also displayed by *Hp*OCBS with substrate OAS and reduced sulfur.

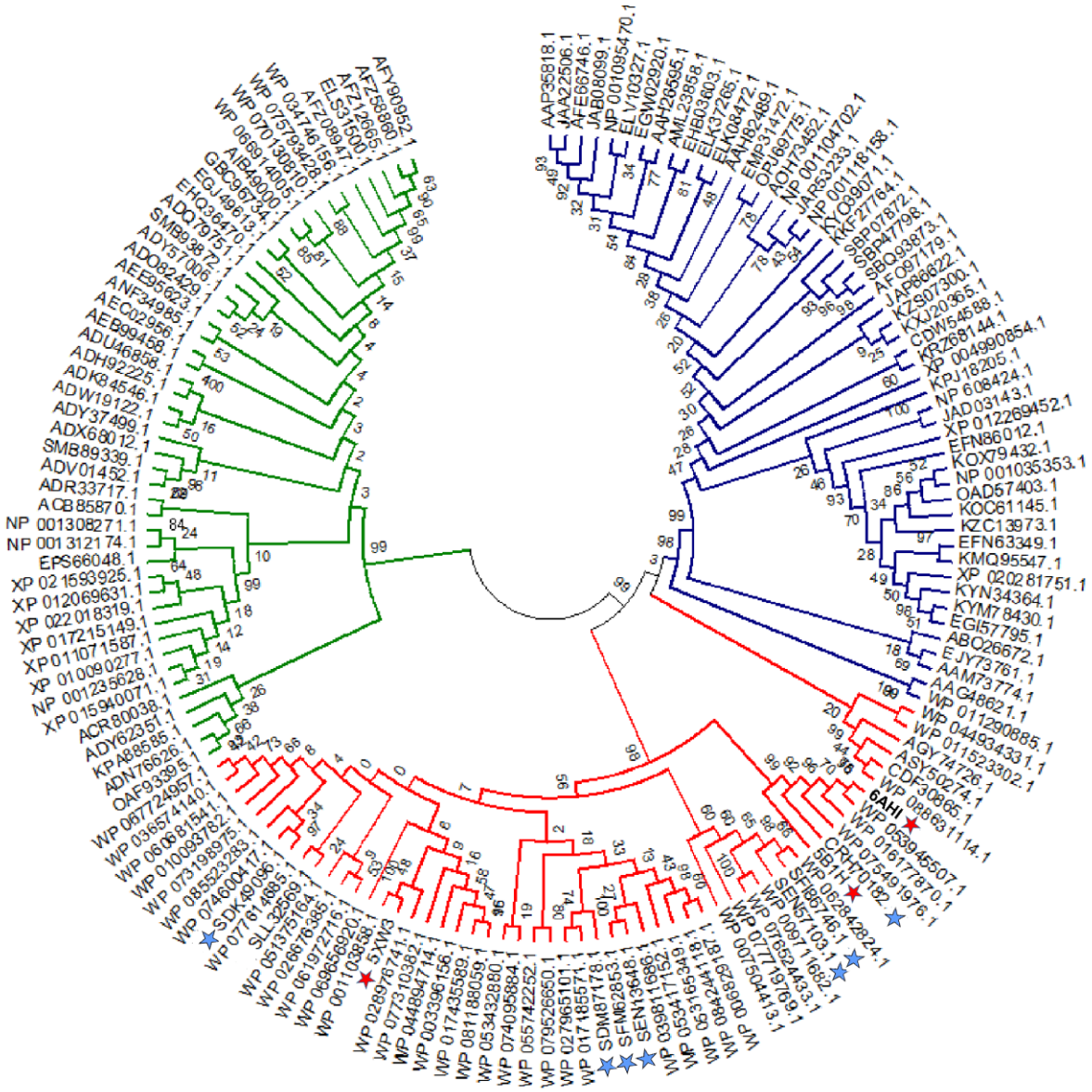
Figure 6. *Hp*OCBS structure. (a) *Hp*OCBS formed a dimer with two protomers interacting with each other in the crystal structure. (b) Each protomer was found to include N- and C-terminal domains consisting of α/β motifs, with these motifs having topologies similar to those in type-II PLP enzymes. α-helices, β-strands and 3_{10} helices are marked accordingly from the N-terminus to the C-terminus. (c) $2F_o$-F_c electron density (at a 1σ cutoff) of PLP covalently bound to Lys46 (LLP46).

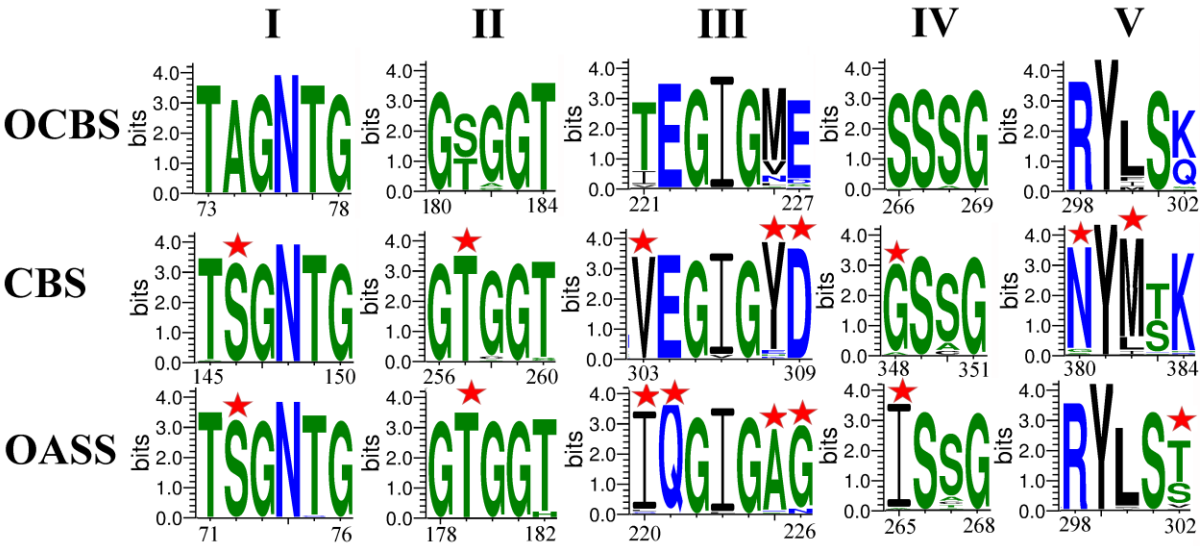
Figure 7. Geometry of the *Hp*OCBS active site with bound methionine. (a) Residues interacting with methionine in the methionine-bound *Hp*OCBS structure. Methionine (blue) was observed to form hydrogen bond (orange) and hydrophobic interactions with surrounding residues (yellow). Lysine-pyridoxal-5`-phosphate (LLP46) is shown in pink. (b) Ligplot showing the residues interacting with methionine (number 601) in chain B. The inset shows $2F_o$-F_c electron density (1σ cutoff) of methionine. (c) Superposition of chain A and chain B (having bound methionine) of *Hp*OCBS revealed the presence of methionine causing a closing of the active site between the N- and C-terminal domains, specifically with abending of the N-terminal domain by 10° leading to a decrease of up to3Å in the distance between the domains. Encircled loops are the asparagine loop and the loop encompassing residues 222-227, both of which apparently moved upon binding of methionine. (d) Superposition of chain A (pink) and chain B (green) showing bound methionine near the PLP at the active site. Binding of methionine resulted in closing of the active site and coincided with the formation of hydrogen bonds by Glu222 with Lys302 of the C-terminal tail and with Ser101 of the N-terminal domain.

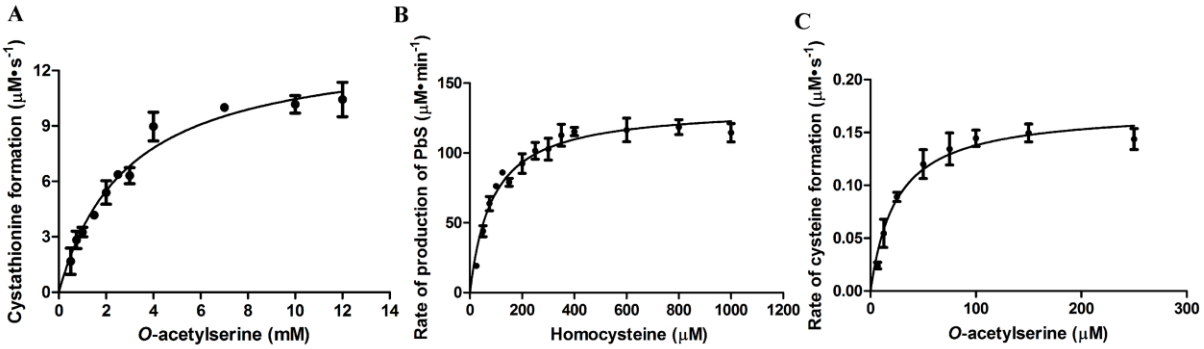
Figure 8. Structural superposition of *Hp*OCBS with OASS and CBS. (a) Structural superposition of *Hp*OCBS (yellow) with OASS (PDB ID: 2Q3B blue) and (b) CBS (PDB ID:1M54 purple) to highlight the differences in residues near the active site. Active site residues of OASS and CBS differing from those of OCBS are shown with dark colours with side chains while only the main chain is shown for identical residues found in different members of the PLP-II family. The five conserved blocks are also labelled. See also Figure S1.

Figure 9. Crucial residues for substrate specificity in OCBS and CBS. (a) Superposition of *H. pylori* OCBS and *S. cerevisiae* CBS structures showing key residues involved in the interaction with the first substrate and thus discriminating between OAS and L-serine. (b) Superposition of active sites of *H. pylori* OCBS and *L. plantarum* OCBS with bound methionine.

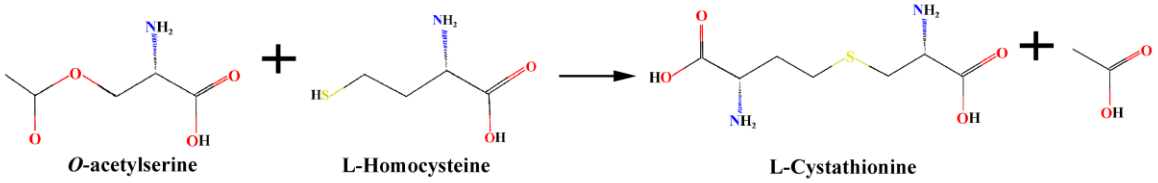






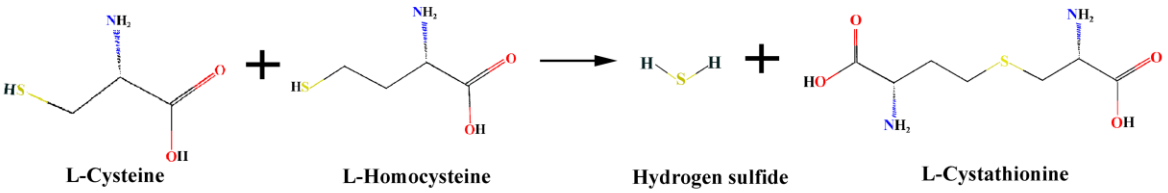


1) *O*-acetylserine-dependent cysthationine β-synthase reaction



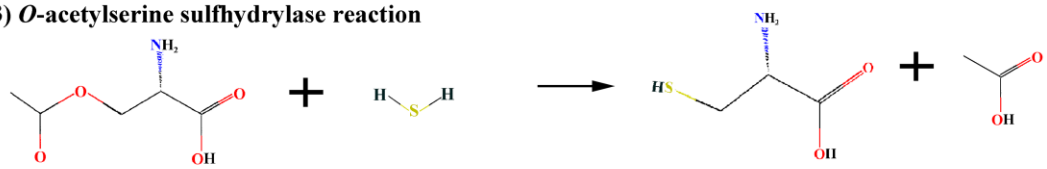
***O*-acetylserine** + **L-Homocysteine** → **L-Cystathionine**

2) Reaction producing hydrogen sulfide



L-Cysteine + **L-Homocysteine** → **Hydrogen sulfide** + **L-Cystathionine**

3) *O*-acetylserine sulfhydrylase reaction



***O*-acetylserine** + **Hydrogen sulfide** → **L-Cysteine**